https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

The Canadian Surface Reanalysis (CaSR) v3.2 precipitation dataset: A 45-year high-resolution analysis for North America (1980–2024)

Dikraa Khedhaouiria [1], Nicolas Gasset [1], Vincent Fortin [1], Milena Dimitrijevic [1], Maxim Bulat [2], and Xihong Wang [2]

[1]Meteorological Research Division, Environment and Climate Change Canada, Dorval, QC, Canada
[2]Meteorological Service of Canada, Environment and Climate Change Canada, Dorval, QC, Canada

Correspondence: Dikraa Khedhaouiria (dikraa.khedhaouiria@ec.gc.ca)

Abstract.

The Canadian Surface Reanalysis (CaSR) includes an offline, high-resolution gridded precipitation reanalysis designed to provide accurate estimates across North America. This product, referred to as CaPA-24h, builds on the Canadian Precipitation Analysis (CaPA) system of Environment and Climate Change Canada (ECCC). It combines a dense network of daily surface observations with a background field from the CaSR dynamical component, using updated quality-control and assimilation procedures to filter spurious observations. This study evaluates the CaPA-24h fields produced in CaSR v3.2, together with their background field, and compares them with the previous version (v2.1) as well as with two independent datasets, ERA5-Land and PRISM. Results show substantial improvements in v3.2, particularly in data-sparse regions, with an enhanced representation of precipitation events of different intensities. Compared to ERA5-Land, CaPA-24h v3.2 provides more accurate seasonal and regional precipitation patterns, while evaluations against PRISM confirm this improved performance. However, biases persist in southern and western mountainous areas, especially for orographic precipitation. A first-time assessment of the hourly disaggregated product reveals limitations in the diurnal cycle representation, indicating the need for refined disaggregation methods and background field generation. Overall, CaPA-24h v3.2 delivers a reliable and well-established gridded precipitation dataset, offering a valuable resource for hydrological, climatological, and impact studies across North America.

1 Introduction

Precipitation is a key component of the Earth's climate system and plays a crucial role in numerous applications, including water resource management, flood forecasting, agriculture, and ecosystem monitoring. Despite its importance, precipitation remains challenging to observe accurately due to its high spatiotemporal variability and the limitations of both direct measurements and indirect estimates (satellites, radars and numerical models; Kidd et al., 2017). As a result, a wide range of precipitation datasets has been developed using diverse sources and methodologies (Beck et al., 2017). However, the abundance of datasets does not eliminate the persistent need for accurate, consistent, and reliable precipitation estimates suitable

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

for scientific and operational use. Furthermore, since each dataset has its own strengths and weaknesses, selecting the most suitable product for a given application requires a thorough characterization and evaluation of its performance, ideally through comparison with independent reference datasets.

Reanalysis has emerged as a valuable tool for climate and meteorological studies, especially in regions with sparse observational networks. By assimilating historical observations into state-of-the-art numerical weather prediction (NWP) models, reanalysis provides physically consistent reconstructions of past atmospheric, surface, and oceanic states (Bengtsson and Shukla, 1988; Bosilovich et al., 2008). These datasets offer broad spatial and temporal coverage, typically spanning several decades, at global or regional scales. However, the quality and usefulness of reanalysis products can vary depending on region, variable, resolution, and intended application (Dee et al., 2011; Gehne et al., 2016; Parker, 2016).

Over the past decades, several major global and regional atmospheric reanalyses have been released, including ERA5 from ECMWF (Hersbach, 2000), JRA-55 from the Japan Meteorological Agency (Kobayashi et al., 2015), and NOAA-NCEP's Climate Forecast System Reanalysis (CFSR) (Kanamitsu et al., 2002), the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2, Gelaro et al., 2017). As demonstrated by their widespread use in the scientific literature, these datasets have become essential for climate research and, more recently, for the development of AI–based forecasting models (Bi et al., 2023). Alongside these global atmospheric products, specialized reanalyses targeting individual Earth system components — such as land, ocean or precipitation — have gained increasing attention (e.g., Zuo et al., 2019; Gasset et al., 2021; Muñoz-Sabater et al., 2021). These products respond to the need for higher spatio-temporal resolution and often rely on large-scale atmospheric reanalyses to provide them with initial and/or boundary conditions. While a comprehensive inventory is beyond the scope of this paper, a review of available reanalyses can be found in Baatz et al. (2021) study. The present work focuses specifically on high-resolution precipitation reanalyses, which are particularly relevant for hydrological applications.

Unlike other variables, precipitation is not directly assimilated in atmospheric reanalyses, but rather produced as an output that drives their land surface components. For instance, some systems ingest observation-based gridded precipitation products, such as CFSR (Saha et al., 2010), while others rely on in-situ station data, as in the Canadian Surface Reanalysis (CaSR) (Gasset et al., 2021; Khedhaouiria et al., 2025). These indirect assimilation strategies can improve the spatial structure and realism of precipitation fields and also influence related surface variables such as soil moisture and ground temperature (Nykanen et al., 2001). However, even in coupled atmosphere–land surface systems, classical precipitation reanalyses are often limited in capturing local extremes (Hu and Franzke, 2020), orographic effects, and the diurnal cycle, all of which are critical for impact studies. Offline approaches that merge NWP backgrounds with station networks help overcome these limitations by delivering datasets more closely tied to observations (Devers et al., 2021).

Examples include CERRA-Land providing 5.5 km daily precipitation fields over Europe by assimilating daily totals from synoptic and climate stations through the SAFRAN system (Verrelle et al., 2022; Soci et al., 2016), and FYRE Climate, delivering daily precipitation and temperature fields back to 1871. A comparable approach is applied in Canada through the offline precipitation reanalysis component of the Canadian Surface Reanalysis (CaSR; Gasset et al., 2021; Khedhaouiria et al., 2025), built upon the Canadian Precipitation Analysis (CaPA) system (Mahfouf et al., 2007; Fortin et al., 2018). In the CaSR

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

configuration, CaPA assimilates 24-hour precipitation totals valid at 12UTC from surface stations across the North American domain, forming the offline component of the reanalysis system.

The recent release of CaSR version 3.2 provides the basis for a new offline precipitation reanalysis. This study documents the configuration of the CaSR v3.2 offline precipitation component (CaPA-24h) and evaluates its performance against CaSR v2.1 and independent reference datasets, with a focus on categorical skill scores, high-impact precipitation metrics, and the diurnal cycle. In addition, the potential use of the operational CaPA precipitation analysis to complement CaSR v3.2 in near-real-time applications is assessed.A comprehensive description and evaluation of the full CaSR v3.2 system will be published in the separate, companion paper. The paper is structured as follows: Sect. 2 describes the assimilation and validation datasets; Sect. 3 presents the analysis methodology; Sect. 4 describes the distributed variables; Sect. 5 details the verification approach; Sect. 6 presents and discusses the results; and Sect. 7 assesses the potential of the operational CaPA for near-real-time use.

2 Datasets

2.1 Precipitation reanalysis fields from CaSR

CaSR consists of two complementary components, each generating precipitation fields distributed to users (Sect. 4). The *online* component provides near-real-time regional surface reanalyses by coupling a Numerical Weather Prediction (NWP) model with a land-surface data assimilation system. In CaSR, the NWP system is the Regional Deterministic Reforecast System (RDRS), a reforecast-based adaptation of the operational RDPS (Caron et al., 2015; Gasset and Milewski, 2024). RDRS runs twice daily at 00 and 12 UTC, producing 48-hour integrations. The land-surface assimilation is performed by the Canadian Land Data Assimilation System (CaLDAS; Carrera et al., 2015), which uses a legacy 6-hourly configuration of the Canadian Precipitation Analysis (hereafter CaPA-6h) to improve the precipitation forcing applied to its land-surface model. It is important to note that CaPA-6h precipitation fields are not distributed directly to users, as they serve primarily as internal forcing for CaLDAS. Instead, users access the output of the coupled RDRS–CaLDAS system, which includes hourly precipitation and related surface variables.

The *offline* component generates an a posteriori daily precipitation reanalysis (hereafter CaPA-24h) using forecasts from the coupled system as background. In CaSR v3.2, the 6–12 h and 12–18 h lead times are combined to form the 24-hour background field (Figure A1, Appendix A), consistent with earlier operational CaPA versions. A limitation of this approach is the reliance on different lead times depending on the hour of the day, which may introduce discontinuities and systematic biases (see Sect. 6.5 and Conclusions). For simplicity, the terms offline precipitation reanalysis, CaPA-24h, and CaSR precipitation reanalysis are used interchangeably hereafter.

Major innovations in CaSR v3.2 are the modernization of its underlying numerical weather prediction (NWP) model and the use of ERA5 (Hersbach et al., 2020) instead of ERA-Interim atmospheric initial and boundary conditions. The system now employs an updated configuration of GEM (Global Environmental Multiscale; Girard et al. 2014), which has been operational since June 2024 (Gasset and Milewski, 2024). This updated configuration includes modernized physical parameterizations, higher vertical resolution, improved topographic representation, and several additional refinements (McTaggart-Cowan et al.,

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

2019). Together, these changes affect the structure and quality of the precipitation background fields and, consequently, the error characteristics of the precipitation analysis relative to earlier CaSR versions.

2.2 Assimilated precipitation from surface stations

The offline precipitation reanalysis based on the CaPA system assimilates precipitation observations from surface stations, primarily sourced from networks operating in Canada and the United States. As shown in the right panel of Figure 1, station density is higher in more populated regions, particularly in southern Canada and the eastern United States. Details on contributing networks, gauge types, and spatial coverage are provided in Table A1 in appendix A.

A key consideration is the change in data sources around the year 2000. Before 2000, observations were obtained from the Integrated Surface Database (ISD; Smith et al. 2011), whereas after 2000, data were sourced from ECCC operational archives, which do not extend as far back in time. In CaSR v3.2, the ISD dataset was fully updated and reprocessed compared to v2.1, taking advantage of improvements made by NCEI (National Centers for Environmental Information) and allowing for refinement of several processing aspects, including the incorporation of trace precipitation information and more accurate station location definitions. This shift of data source explains the sharp discontinuity in the number of assimilated stations (Figure 1, left panel), increasing from an average of about 2000 stations before 2000 per analysis to approximately 9000 afterward. The use of ISD data prior to 2000 also introduced additional challenges. The ISD format and metadata structure differ substantially from those of ECCC archives and from the standardized input expected by CaPA, requiring extensive preprocessing to harmonize station identifiers, timestamps, and precipitation codes, and to reconstruct missing metadata such as gauge type. In addition, the spatial coverage and data completeness of ISD records vary across networks and time periods, adding further uncertainty when matching ISD stations with their modern ECCC counterparts.

All precipitation observations undergo automated quality control (QC) procedures consistent with those applied in the operational CaPA system (Lespinas et al., 2015; Khedhaouiria et al., 2022). These procedures, applied at each assimilation time step, include spatial and temporal consistency checks such as leave-one-out validation with neighboring stations. Additional filters account for seasonal effects, including intense summer storms and wintertime windy conditions that can affect measurement reliability. Seasonal variability in the number of assimilated 24h accumulation observations is also evident in Figure 1. Although designed for real-time operations, these QC methods have proven effective in the reanalysis context, but with some limitations. In particular, persistently low precipitation values from certain networks were not flagged by the automated QC. These issues were revealed during an intermediate CaPA-24h run, where monthly accumulations showed that some stations consistently reported unrealistically low totals over specific periods. Such stations were subsequently excluded from the assimilation. The same intermediate run also helped identify suspicious extreme events that escaped initial QC screening. Specifically, stations reporting monthly precipitation exceeding 300 mm while also exceeding five times the corresponding background-field total, were flagged and removed. This additional diagnostic step thus provided a complementary layer of quality assurance beyond the standard real-time QC procedures.

Finally, two Canadian datasets, Adjusted Daily Rainfall and Snowfall (AdjDlyRS) and Adjusted Hourly Rain and Snow (AdjHlyRS, newly included in CaSR v3.2), receive special treatment in the QC workflow. Unlike other data sources, these

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

datasets are exempt from standard CaPA quality control filters because they have been pre-processed to correct known observational biases. AdjDlyRS addresses errors in manual gauge measurements, including undercatch, evaporation, and wetting losses (Wang et al., 2017), while AdjHlyRS provides bias-adjusted hourly precipitation from automated stations (Smith et al., 2022). Since these corrections already account for conditions that would normally trigger rejection (e.g., solid precipitation or high wind), applying the standard CaPA QC filters would be inappropriate and could discard valid bias-corrected observations. However, their inclusion introduces a potential for duplicate records, a known issue in reanalysis assimilation (Bell et al., 2021), since they originate from the same underlying SYNOP and ECCC sources already ingested separately. To prevent redundancy, the reanalysis retains only the adjusted record when an AdjDlyRS or AdjHlyRS station is located within 0.02° of an existing SYNOP or ECCC record, giving preference to the bias-corrected dataset. Although excluded from some rejection filters process, these adjusted observations are still subject to the spatial consistency checks implemented within the CaPA quality-control framework (Lespinas et al., 2015).

The integration of additional datasets, such as AdjHlyRS in CaSR v3.2, led to a substantial increase in the number of assimilated observations after 2000. This increase is clearly reflected in the differences between v3.2 and v2.1 (Figure 1, subplot of the differences). While ISD provided valuable information prior to 2000, it is an external dataset over which there is limited control. In contrast, the enhanced data availability in v3.2 represents a tangible improvement in both the density and quality of assimilated observations in the offline precipitation reanalysis.

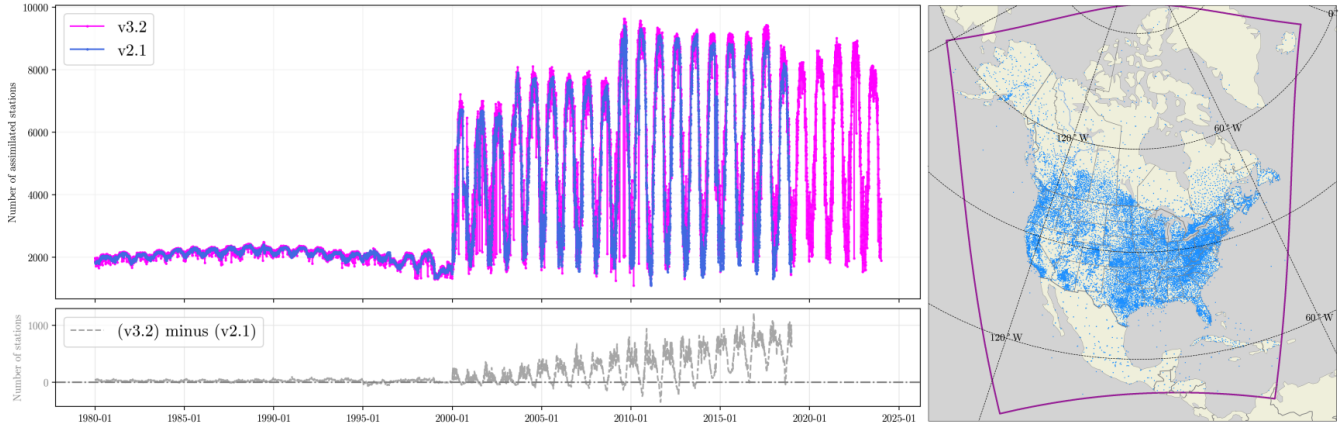

Figure 1. *Left*: Daily number of surface stations assimilated in the offline precipitation reanalysis from CaSR v2.1 (blue) and v3.2 (magenta), smoothed using a 7-day running mean. The bottom subplot shows the difference in station counts between v3.2 and v2.1. *Right*: Spatial distribution of assimilated stations, aggregated on a 0.05° grid. The CaSR reanalysis domain is outlined in purple.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

2.3 Reference gridded datasets for evaluation

2.3.1 ERA5-Land

Precipitation fields from ERA5-Land (Muñoz-Sabater et al., 2021) – the land component of ECMWF's fifth-generation reanalysis – were selected for comparison with the CaSR offline precipitation reanalysis. Both ERA5-Land and CaSR v3.2 share a common atmospheric foundation: they are both informed by atmospheric data from the ERA5 reanalysis (Hersbach et al., 2020), ensuring consistency in large-scale atmospheric conditions. Their purposes, however, differ. ERA5-Land is produced by running an offline global land surface model driven by ERA5 atmospheric fields, without direct assimilation of precipitation observations. In contrast, CaSR online component is initialized with ERA5 upper air fields but generates its own precipitation forecasts through a regional atmospheric model, while its offline precipitation component explicitly assimilates surface observations to reconstruct high-resolution precipitation fields.

ERA5-Land provides hourly surface variables at a horizontal resolution of approximately 9 km, comparable to that of CaSR, which facilitates scale-consistent comparisons. For this study, hourly ERA5-Land precipitation data spanning 1980–2024 were retrieved from the Copernicus Climate Change Service (CDS, Copernicus) and interpolated onto the CaSR grid using a conservative mapping approach (Jiawei Zhuang et al., 2025). Although ERA5 assimilates radar-derived precipitation from Stage IV over the United States (Lopez, 2011), documented precipitation biases in ERA5 can propagate into ERA5-Land (Cucchi et al., 2020; Muñoz-Sabater et al., 2021).

The objective of this comparison is to assess how CaSR v3.2 offline component product compares relative to ERA5-Land within the North American domain. To ensure a fair assessment, precipitation fields from the CaSR online component – produced without direct assimilation of precipitation observations are also included in the verification – providing a baseline that isolates the added value of the offline assimilation.

2.3.2 PRISM dataset

PRISM dataset (Parameter-elevation Regressions on Independent Slopes Model; Daly et al. 2021, 2008) is employed to evaluate the precipitation reanalysis. Developed by the PRISM Climate Group at Oregon State University for the contiguous United States of America (CONUS), PRISM provides observation-based, gridded climate data at multiple spatial and temporal resolutions, including daily and monthly precipitation. Precipitation at each grid cell is estimated by combining nearby station observations with several geographic predictor grids, with greater weight given to stations that are close to the target grid cell, at similar elevations, and located on comparable slopes or terrain features. A comprehensive description of the model algorithms and input data is provided by Daly et al. (2008).

For the present study, daily precipitation fields covering 1981–2023 at a horizontal resolution of $\Delta x \simeq 4$ km were initially considered for comparison with CaSR. An intermediate analysis conducted over the 1981–2018 period revealed temporal discontinuities in PRISM-derived precipitation statistics, particularly during summer. These discontinuities are likely associated with changes in the underlying observing system and processing methodology, including the progressive incorporation of radar-based information into PRISM (Daly et al., 2021). To minimize the influence of such inhomogeneities on the inter-

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

comparison, the PRISM dataset was therefore restricted to the 2002–2018 period, for which the daily fields exhibit improved temporal consistency.

An alternative PRISM product designed for long-term climate analyses is also available; however, this dataset is provided at monthly temporal resolution only, which is less suitable for the present study focusing on daily precipitation characteristics, frequency–intensity decomposition, and high-impact precipitation metrics. Consequently, daily PRISM fields over the 2002–2018 period were retained and upscaled to the CaSR grid ($\Delta x \simeq 10$ km) using conservative interpolation approaches (Jiawei Zhuang et al., 2025).

3 Offline precipitation reanalysis methodology

The CaPA system used in the offline reanalysis assimilates precipitation observations from multiple sources using an optimal interpolation (OI) algorithm, consistent with the most recent operational implementation of CaPA (Lespinas et al., 2024). The methodology underlying this approach has been extensively documented in previous studies (Khedhaouiria et al., 2022; Fortin et al., 2018; Mahfouf et al., 2007) and is not repeated in detail here. However, it is worth recalling that, as with most data assimilation systems, the analysis at each grid point is governed by the specification of two key components: the observation error covariance matrix R and the background error covariance matrix B.

The observation error covariance matrix R is assumed to be diagonal for surface stations, reflecting the hypothesis of uncorrelated observation errors. Its diagonal elements represent the squared standard deviation of the observation errors (σ_o^2). In contrast, the background error covariance matrix B accounts for spatial correlations in the model background field and is parameterized as:

$$B_{i,j} = \sigma_b^2 \exp\left(-\frac{d_{i,j}}{l_b}\right), \tag{1}$$

where σ_b denotes the standard deviation of the background errors, l_b is the correlation length scale, and $d_{i,j}$ is the distance between locations i and j.

A distinctive feature of CaPA is that, unlike other systems where σ_o, σ_b, and l_b are fixed, these parameters are dynamically estimated using variographic analysis of the innovations over the domain in a preliminary step of the analysis (Fortin et al., 2015). This data-driven estimation allows the parameters to evolve daily according to the prevailing meteorological conditions. A limitation, however, is that the resulting error parameters remain spatially uniform within the domain on a given analysis time step.

Figure 2 shows the annual cycle of σ_o and σ_b for the CaSR v3.2 offline precipitation reanalysis. Both errors exhibit clear seasonal variability, with larger values in summer and autumn and lower values in winter and spring, consistent with previous studies (Desroziers et al., 2005; L'Ecuyer and Stephens, 2002). Reported values in Figure 2 are estimated in the Box–Cox-transformed space required by the OI algorithm (Fortin et al., 2015), which prevents direct interpretation in precipitation units. Finally, the estimated σ_o^2 values show greater variability due to the lower density of observation stations during the pre-2000 period, which results in noisier estimates.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

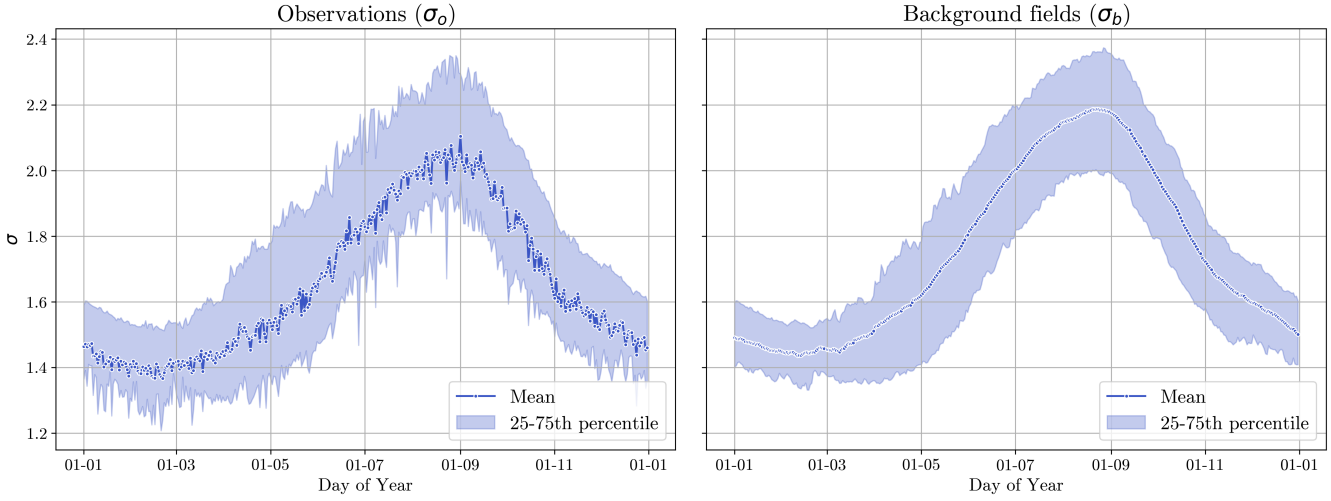

Figure 2. Mean daily standard deviation of observation errors (σ_o, left) and background errors (σ_b, right) in the CaPA Box-Cox transformed space over the CaSR v3.2 domain, averaged over the 1980–2023 period. Shaded areas represent the interquartile range (25th–75th percentile), illustrating the inter-annual variability.

4 Output variables distributed to users

Four types of precipitation-related variables are publicly available from CaSR. Hourly precipitation from the online component provides the short-range forecasts that serve as background fields in the assimilation process. The offline reanalysis product corresponds to the daily 24-hour accumulated precipitation field generated by CaPA, valid at 12 UTC. A complementary variable, the Confidence Index of the Analysis (CFIA), is also provided at daily resolution; it ranges from 0 to 1 and quantifies the degree to which each grid point is constrained by observations (see Fortin et al., 2015). Finally, hourly reanalysis precipitation is derived by temporally disaggregating the daily offline reanalysis, as detailed below.

The disaggregation follows a two-step linear procedure, consistent with previous CaSR versions (Section 2.3 in Gasset et al., 2021). The approach transforms daily precipitation totals into hourly values by making use of short-range forecasts of the online component. In the first step, 6-hourly CaPA analyses (CaPA-6h) are used to adjust the timing of the RDRS hourly forecast fields so that their 6-hour accumulations match the observations (Step 1 in Figure 3). In the second step, the adjusted hourly values are rescaled to ensure that their 24-hour sum matches the CaPA-24h analysis, which serves as the final daily constraint (Step 2 in Figure 3). This procedure preserves the realistic temporal distribution of precipitation provided by the online component while maintaining full consistency with the 6-hour and 24-hour totals generated by CaPA.

When the CaSR online component indicates no precipitation within a 6- or 24-hour window but the corresponding CaPA analysis reports a nonzero accumulation, the total CaPA precipitation is evenly redistributed across the 6 or 24 hours. While this enforces temporal consistency, it can also introduce artefacts, such as sequences of constant hourly precipitation when online forecasts are zero. These artefacts, often localized near observational stations due to the influence of optimal interpolation, lead

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

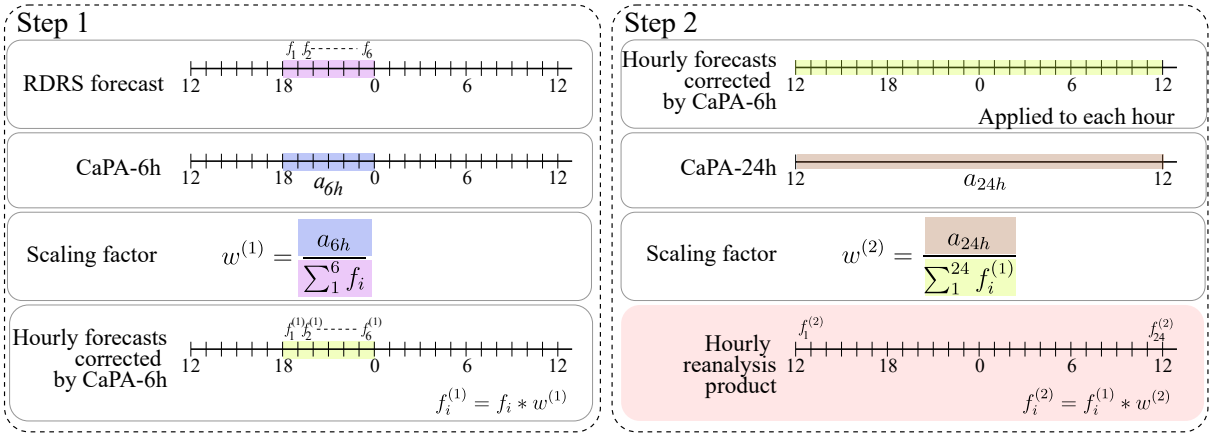

Figure 3. Step 1 (left) illustrates the temporal disaggregation of CaPA-6h using RDRS hourly forecasts. Step 2 (right) shows how the CaPA-24h analysis is further disaggregated using the Step 1 output to generate the hourly reanalysis distributed to the public. Here, a_{6h} and a_{24h} denote the 6-hour and 24-hour CaPA analyses, respectively, and f_i represents the RDRS hourly forecasts.

to an overrepresentation of event durations that are exact multiples of 6 or 24 hours. This distortion in the distribution of event durations may in turn affect derived hydrological indicators. Potential improvements for future CaSR versions include more advanced disaggregation approaches based on data-driven temporal profiles (including machine learning methods), stochastic perturbations, or hybrid techniques.

Finally, CaSR v3.2 variables are available for 1980–2024. To ensure temporal continuity with near-real-time precipitation data, users may refer to the near real time operational CaPA product to fill gaps. However, the operational CaPA differs from the reanalysis in several aspects, and the potential impacts of using it are detailed in Section 7.

5 Evaluation framework

5.1 Direct comparison with in situ gauges

The verification process relied on the Frequency Bias Index (FBI) and the Equitable Threat Score (ETS), both derived from a 2×2 contingency table for binary precipitation events, and on the partial mean to characterize conditional precipitation intensity. The FBI measures the ratio between the frequency of predicted and observed events, while the ETS quantifies the fraction of correctly predicted events after accounting for hits expected by random chance. The partial mean provides complementary information by describing the average precipitation intensity conditional on event occurrence below a given threshold, thereby allowing frequency- and intensity-related errors to be disentangled. Definitions of metrics are provided in Appendix D. Precipitation events were classified as “1” when daily accumulation exceeded a given threshold and “0” otherwise. These thresholds corresponded either to absolute values (ranging from 0.2 to 100 mm) when the entire domain was considered, or to percentiles (0, 20, 50, 70, 80, 95, and 99.9) computed from the non-zero observed precipitation distribution when regional analyses were

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

performed. A minimum threshold of 0.2 mm was applied to define non-zero 24-hour precipitation, both to remove drizzle-level background noise and to match the detection limits of standard precipitation gauges (typically 0.1–0.3 mm).

All metrics were computed for the CaSR offline reanalysis (versions v3.2 and v2.1), its background fields, and ERA5-Land precipitation. The reanalysis estimates used in the verification were derived using the leave-one-out (LOO) framework embedded within the CaPA algorithm. This approach mitigates the artificial skill inflation that would otherwise arise from using the same station observations in both assimilation and verification. Specifically, CaPA generates LOO precipitation analyses at each station location by sequentially removing that station from the assimilation process and estimating precipitation there based solely on surrounding observations. Although this approach does not fully eliminate spatial autocorrelation between validation points (Panthou et al., 2012), it provides an operationally feasible method for independent verification.

Because the in situ network is highly heterogeneous in both space and time, a substantial thinning and quality-selection procedure was applied before verification to ensure spatial representativeness and observational reliability across the domain. Verification relied only on trusted gauges: synoptic stations during the warm season (JJA and SON), and manual synoptic stations during the cold and transition seasons (DJF and MAM), which are better suited for measuring solid precipitation. In addition, adjusted precipitation stations from the Canadian network (AdjDlyRS and AdjHlyRS) were included to enhance the reliability of the verification.

Spatial thinning was performed to reduce over-representation of densely instrumented regions, such as southern Canada and the northeastern United States, by retaining at most one station within each 0.1°×0.1° grid cell. Within each cell, the station reporting the largest number of valid daily observations over the evaluation period was retained. Temporal thinning was also applied by requiring a minimum data availability of 8% over the evaluation window. Although this threshold may appear low, it was selected after testing several values and found to be the minimum level at which verification scores stabilized, thereby maximizing the retained sample size while preserving score robustness.

Comparisons between CaSR versions 2.1 and 3.2 were performed only for stations whose coordinates and observed precipitation values matched exactly in both datasets. Despite extensive efforts to harmonize station identifiers and perform spatial matching between the two reanalysis versions, a small fraction of stations could not be consistently paired. Retaining an 8% temporal completeness threshold thus helps preserve as many stations as possible for meaningful intercomparison while maintaining statistical robustness. The evaluation was conducted separately for each of the four seasons, over the full 1980–2018 period, as well as for two subperiods (1980–1999 and 2000–2018). This partitioning enables assessment of the potential influence of the reduced number of assimilated stations before 2000 on verification results (Sect 6.2). Maps illustrating the evaluation network after the spatio-temporal thinning are provided in the supplementary material.

Given the large spatial extent of the domain and the diversity of regional climates, the frequency of precipitation events – and consequently the skill scores – can vary substantially across space, complicating interpretation. To account for this variability, all metrics were computed both over the full domain and over predefined subregions. This regional breakdown is particularly important for scores such as the FBI and ETS, which are sensitive to event frequency and thus can be biased by climate regime (Hamill and Whitaker, 2006). The subregions follow the classification of Bukovsky (2011), illustrated in Appendix B.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

5.2 Gridded product intercomparison with alternative datasets

CaPA-24h precipitation reanalysis was compared with ERA5-Land and PRISM from three complementary perspectives: (i) representation of precipitation accumulation, (ii) representation of high-impact precipitation, and (iii) representation of the diurnal cycle of precipitation (the latter restricted to ERA5-Land and CaSR, as PRISM is not available at hourly resolution). Together, these evaluations provide a comprehensive view of how CaSR v3.2 reproduces both the mean and extreme characteristics of precipitation relative to established gridded datasets.

For the first aspect, seasonal precipitation totals were accumulated into 30-year time series (1980–2018) for each dataset and each grid cell, all interpolated to the CaSR v3.2 grid. Maps of mean seasonal accumulations were examined alongside relative differences between CaSR v3.2 and v2.1, CaSR v3.2 and ERA5-Land, CaSR v3.2 and PRISM, and ERA5-Land and PRISM. These comparisons highlight the regional distribution of differences between reanalysis-based datasets (CaSR v3.2, v2.1, ERA5-Land) and the observation-based PRISM dataset. Seasonal time series of accumulated precipitation were also averaged over Bukovsky regions to assess long-term temporal consistency and potential drifts. To better understand the origin of biases in seasonal precipitation totals, the total bias relative to PRISM was further decomposed into contributions from the frequency of wet days (precipitation > 1 mm) and the mean precipitation intensity on wet days. This decomposition allows the identification of regions where small total biases may result from compensating errors between precipitation occurrence and intensity.

Beyond mean accumulation, the ability to reproduce extremes is critical for hydrological and climate applications. High-impact precipitation was therefore assessed using indices from the Expert Team on Climate Change Detection and Indices (ETCCDI; Karl et al., 1999). Four indices were considered: Rx1day (annual maximum 1-day precipitation), Rx5day (annual maximum 5-day precipitation), R95pTOT (annual precipitation from days exceeding the 95th percentile), and R99pTOT (same as R95pTOT but with the 99th percentile). Since results for Rx5day and R99pTOT closely resemble those of Rx1day and R95pTOT, respectively, only Rx1day and R95pTOT are presented here. Time series of indices were compared across seasons and for each grid cell using the Kling–Gupta Efficiency score (KGE; Kling et al., 2012; Gupta et al., 2009) and its components:

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\alpha-1)^2} \tag{2}$$

$$\beta = \frac{\mu_{reanalysis}}{\mu_{PRISM}} \tag{3}$$

$$\alpha = \frac{\sigma_{reanalysis}}{\sigma_{PRISM}} \tag{4}$$

where r is the Pearson correlation coefficient, β the mean bias ratio (ratio of means, μ), and α the variability ratio (ratio of standard deviations, σ). A KGE of 1 indicates perfect agreement with PRISM, while deviations in α, β, or r highlight whether errors are dominated by bias, variability, or correlation. All indices were computed at the grid-cell level after interpolating each dataset to the CaSR v3.2 grid to ensure spatial consistency.

Finally, the diurnal cycle of hourly precipitation was examined, as it is a widely used diagnostic of model performance and physical process representation (Dai et al., 1999). This evaluation is particularly relevant given that both offline and online

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

precipitation fields for a given day are constructed using forecast segments of varying lead times (see Section 2.1). It also provides insights on the hourly fields distributed to users (Section 4), which are less frequently assessed than daily totals. The diurnal cycle was computed for each grid cell from hourly precipitation in the CaSR v3.2 offline reanalysis and ERA5-Land, averaged over 1980–2018 and stratified by season and region. For each region, the median cycle and interquartile range are presented. Although a direct observational benchmark is not included in this study, this analysis remains valuable for identifying structural differences in the representation of sub-daily precipitation variability.

6 Results and discussions

6.1 Impact of version upgrade: v3.2 vs. v2.1 precipitation

Figure 4 presents the evaluation of daily precipitation performance over the 1980–2018 period across the domain, using three complementary metrics—FBI, ETS, and the partial mean—computed at multiple precipitation thresholds and for each season.

Overall, differences between CaPA-24h versions 2.1 and 3.2 are relatively small. This is expected, as both versions assimilate largely similar precipitation observations, which naturally leads to comparable analysis statistics at station locations. This interpretation is further supported by the comparison with the corresponding background fields, for which more pronounced differences emerge: the CaSR v3.2 background exhibits reduced bias (FBI values closer to unity and partial means closer to observations) and higher skill (larger ETS values) than v2.1. Both CaSR versions consistently show reduced bias and improved skill relative to their respective background fields, highlighting the added value of the assimilation procedure.

These results are broadly consistent with those reported by Gasset et al. (2021) (see their Figure 9), despite differences in both the evaluation period (2010–2014 versus 1980–2018 here) and the evaluation network. In both studies, the analyses exhibit lower frequency bias (FBI) and higher skill (ETS) than the background fields. A notable difference emerges in winter (DJF) frequency bias at high thresholds, where the longer 1980–2018 period enables improved sampling of rare, high-impact events compared to 2010–2014, leading to a distinct high-threshold signature. This feature is specific to winter conditions, as seasons dominated by liquid precipitation benefit from a much denser station network, whereas winter precipitation is evaluated using a substantially sparser observational network, requiring longer periods to adequately sample intense events. In particular, events exceeding 100 mm day^{-1} occur much more frequently in the background fields (FBI close to 2) than in the analyses (FBI of approximately 1.0–1.1). However, the impact on the partial mean—defined here as the conditional mean precipitation intensity below a given threshold—remains limited, indicating that the excess detections in the background fields are primarily concentrated near the threshold rather than associated with strongly overestimated intensities. This characteristic is consistent across seasons: background fields generally exhibit larger partial means than observations at high thresholds (especially in JJA), while the analyses tend to slightly underestimate the partial mean for events above approximately 25 mm day^{-1}, with overall biases reduced in v3.2 relative to v2.1.

Including ERA5-Land in the comparison reveals a markedly different response. ERA5-Land tends to overestimate the frequency of light to moderate precipitation events while underestimating the occurrence of high-intensity events, resulting in a threshold-dependent performance of the frequency bias. The partial mean further shows that ERA5-Land systematically

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

produces intensities higher than observed at low to moderate thresholds, with values generally comparable to the CaSR background fields up to about 25 mm/day. Beyond this threshold, the partial mean reaches a plateau, indicating an underestimated contribution from higher-intensity events, consistent with the pronounced high-threshold FBI deficit—particularly in summer (JJA). ERA5-Land also exhibits substantially lower skill and larger biases than both CaSR analyses (v2.1 and v3.2) and their respective background fields, as reflected by the ETS metric, in agreement with previous evaluations reported by Muñoz-Sabater et al. (2021).

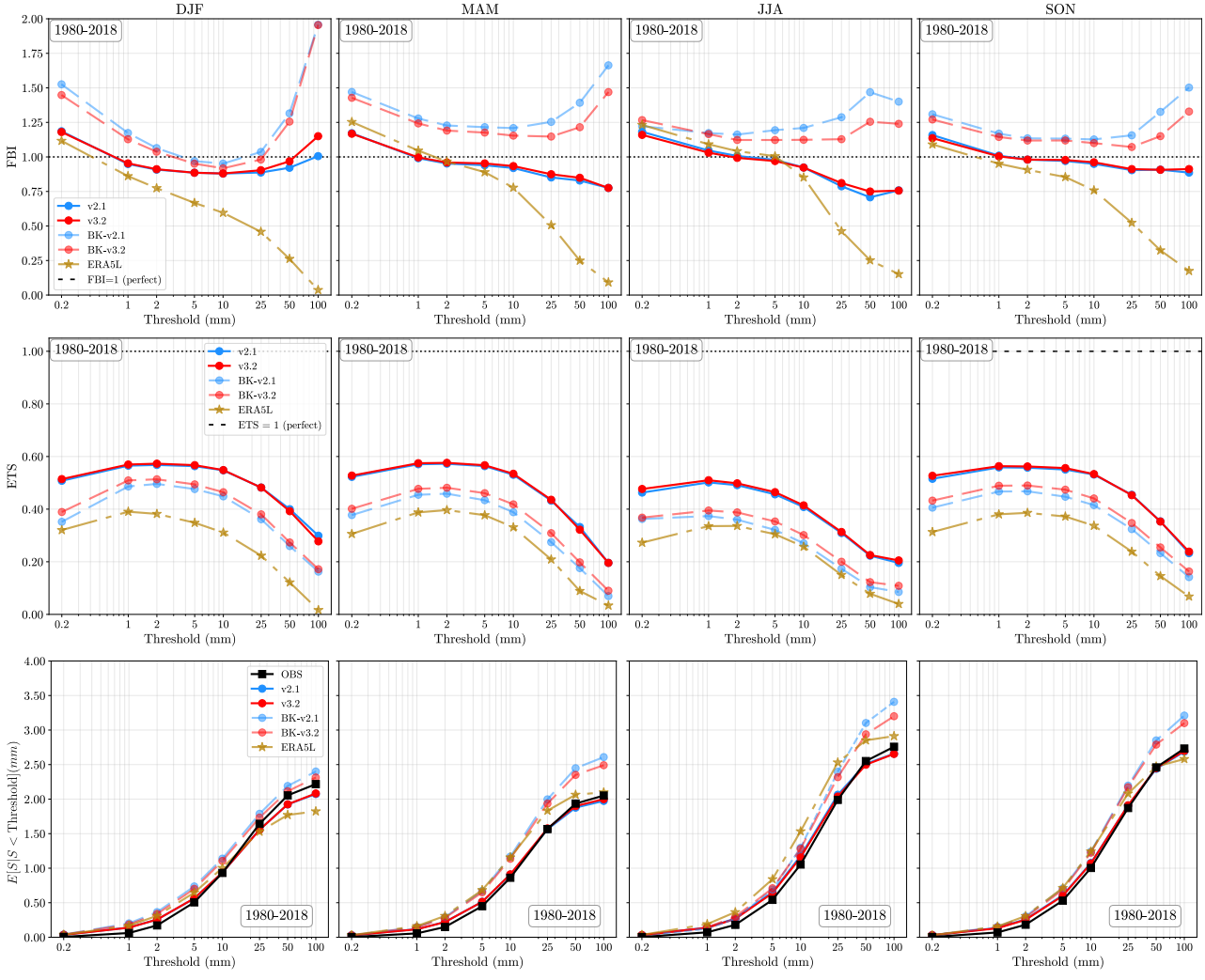

Figure 4. Performance of daily precipitation estimates over the 1980–2018 period. The first row shows the Frequency Bias Index (FBI), the second row the Equitable Threat Score (ETS), and the third row the partial mean $E[S \mid S < \text{Threshold}]$, all shown as a function of the precipitation threshold (mm). Columns correspond to the four seasons (DJF, MAM, JJA, and SON). Results are shown for CaPA-24h v2.1 (blue circles), CaPA-24h v3.2 (red circles), their respective background fields (BK-v2.1 and BK-v3.2; dashed lines in matching colours), and ERA5-Land (gold stars).

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

While domain-averaged scores provide a useful first-order assessment of system performance, they can mask substantial regional differences, particularly in regions characterized by sparse station coverage where the background field exerts a stronger influence on the analysis. Differences between CaPA-24h v3.2 and v2.1 are generally small across ETS, FBI, and the partial mean, indicating comparable skill over most regions, as documented in the Supplemental Material. Noticeable discrepancies are mainly confined to the Desert and South regions, where the background field exhibits an excessive precipitation bias in v2.1 (Gasset et al., 2021), which remains amplified in v3.2. Although the analysis step partially mitigates this degradation through observation assimilation, it does not fully compensate for the background bias, resulting in degraded FBI and partial-mean values, while event-detection skill remains largely unchanged.

More pronounced differences emerge when focusing specifically on the background field (Figure 5), which highlights improvements (degradations) in FBI or ETS scores in red (blue) across regions and seasons. These results indicate an overall reduction in frequency bias in v3.2, while changes in ETS are largely neutral, suggesting that the improvements primarily affect precipitation bias rather than event-detection skill. Winter performance remains more mixed, with a persistent degradation in the southern domain. Overall, these results for the background field are reassuring, as they indicate that in regions with sparse or no station coverage—where the background field dominates in CaSR–CaPA—v3.2 generally provides less biased and more reliable precipitation estimates than v2.1.

6.2 Influence of changes in assimilation data sources on CaPA-24h

Figure 6 presents the FBI and ETS scores for CaPA-24h v3.2, its background field (BK), CaPA-24h v2.1, and ERA5-Land, evaluated over the CaSR domain for winter (DJF) and summer (JJA). Results are shown separately for two periods: 1980–1999, when observations were primarily sourced from the Integrated Surface Database (ISD), and 2000–2018, when ECCC operational archives were used.

Comparisons between the two evaluation periods reveal only limited changes in relative performance and a stable ranking of the datasets. ETS values are generally slightly higher during the 2000–2018 period than during 1980–1999 for most products in DJF at low to moderate thresholds, whereas no systematic differences are observed in JJA, suggesting a modest overall improvement in categorical precipitation skill. The skill gap between background fields and analyses remains comparable across periods, indicating that the relative impact of the assimilation step is maintained despite major changes in the observing system. As in the previous section, ERA5-Land consistently exhibits the weakest performance across all configurations, with lower ETS values and systematically low FBI values indicating an underestimation of precipitation occurrence, particularly at higher thresholds. This feature, also evident in the partial mean, is consistent across seasons and evaluation periods.

Differences between the two periods mainly emerge at moderately extreme thresholds and are primarily reflected in the FBI. These differences can be attributed to sampling effects arising from (i) differences in station density between the two periods, with fewer available stations prior to 2000 by construction, and (ii) a seasonally varying sampling effect, as winter precipitation is evaluated using a sparser station network than liquid-precipitation seasons. This interpretation is further supported by the ERA5-Land results, which also exhibit small but noticeable differences between the two periods despite being unaffected by

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

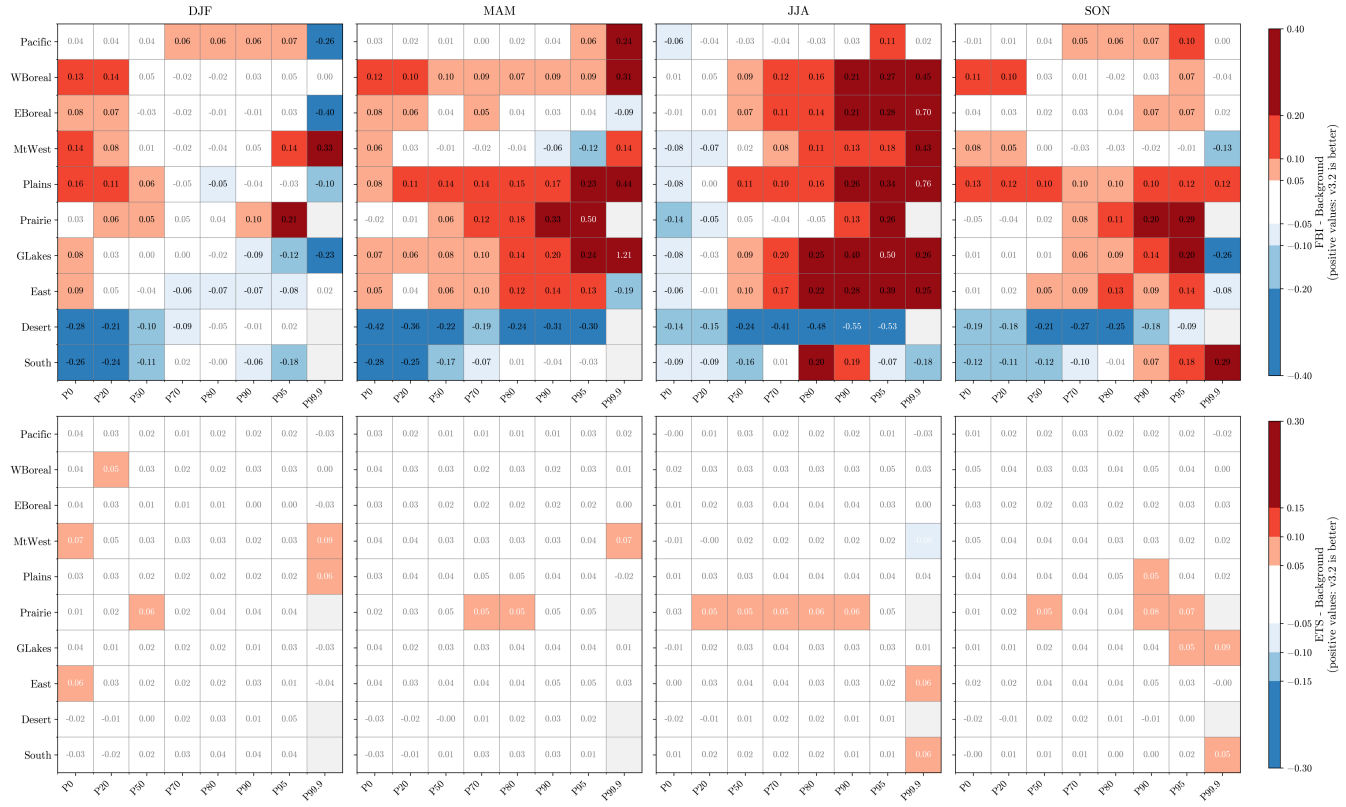

Figure 5. Improvement (red) or degradation (blue) in the Frequency Bias Index (FBI, top row) and Equitable Threat Score (ETS, bottom row) of the precipitation background field from CaSR v3.2 relative to CaSR v2.1, across Bukovsky regions and percentile-based precipitation thresholds. The top panels show $|FBI_{v2.1} - 1| - |FBI_{v3.2} - 1|$, where positive values indicate a reduction in the frequency bias in v3.2, while the bottom panels show $ETS_{v3.2} - ETS_{v2.1}$. Each column corresponds to a season (DJF, MAM, JJA, SON). Gray cells indicate insufficient data to compute the scores.

the change in assimilated precipitation datasets after 2000, indicating that sampling effects may contribute to the observed variations, independently of changes in the observing system.

6.3 Seasonal and regional precipitation accumulation patterns across datasets

Figure 7 shows the seasonal spatial distribution of precipitation and the effects of updates in CaPA-24h v3.2 compared to v2.1. ERA5-Land is included for comparison due to extended use in the community and its matching horizontal resolution with CaSR, and PRISM serves as a high-resolution observational reference.

Comparisons between v3.2 and v2.1 (second row of Fig. 7) show widespread precipitation increases across the southern part of the domain in all seasons, particularly over the Gulf of California and along the Gulf of Mexico, likely associated with changes in the background field and model physics affecting moisture transport and convective processes. These southern

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

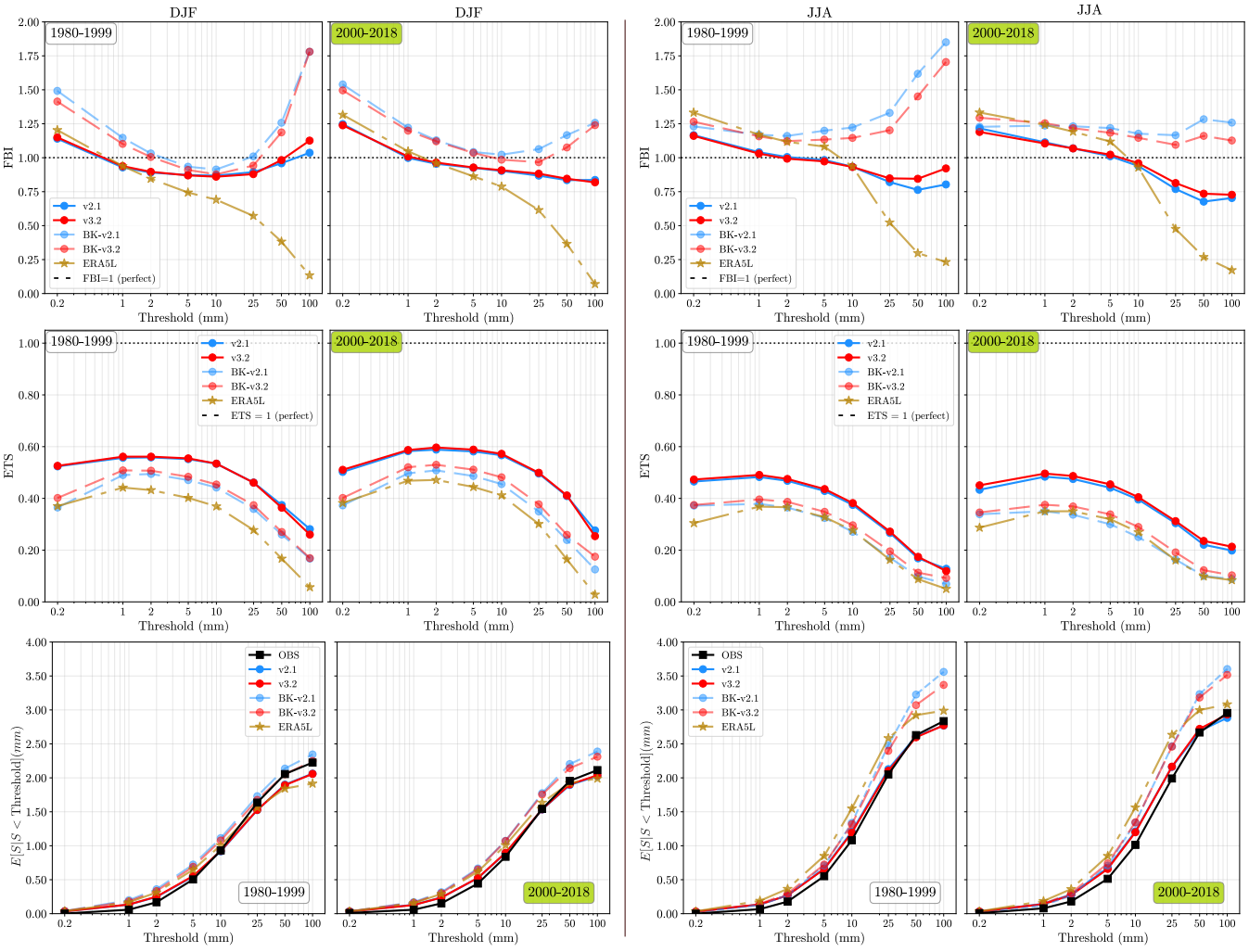

Figure 6. Performance metrics for daily precipitation estimates as a function of precipitation threshold. Results are shown for winter (DJF; left) and summer (JJA; right) and for two evaluation periods: 1980–1999 and 2000–2018. The top row shows the Frequency Bias Index (FBI), the middle row the Equitable Threat Score (ETS), and the bottom row the partial mean. Curves are shown for CaPA-24h v2.1 (blue circles) and v3.2 (red circles), their respective background fields (BK-v2.1 and BK-v3.2; dashed lines in matching colours), and ERA5-Land (gold stars). Dotted horizontal lines indicate perfect scores (FBI = 1, ETS = 1), and the black curve in the partial-mean panels represents observations (OBS).

increases contribute to a degradation of the frequency bias in v3.2 relative to v2.1, especially for small to moderate intensity precipitation events, through an increased false alarm rate in arid regions (e.g., south of the Desert region following the Bukovsky classification). In the Rocky Mountains, precipitation increases especially at high elevations, primarily due to modifications in the CaPA background field linked to GEM model upgrades and a sharper representation of topography (see Khedhaouiria et al., 2025). In summer, precipitation over much of Canada is reduced in v3.2, while regions with dense station coverage (e.g.,

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

along the U.S.–Canada border and in the eastern United States) remain largely unchanged, reflecting stronger observational constraints.The reduced summer precipitation accumulations in northern regions improve skill and reduce bias relative to v2.1, indicating that v2.1 was overall too wet in these areas.

The relative differences between CaPA-24h v3.2 and ERA5-Land exhibit their strongest amplitudes over complex terrain, with consistent structures along the western cordillera in all seasons (Fig. 7). In DJF and SON, a marked sign alternation is apparent across the coastal mountains and the Rocky Mountains: CaPA-24h tends to be drier than ERA5-Land along parts of the Pacific coastal ranges, while becoming wetter over portions of the interior high terrain, consistent with differing representations of orographic precipitation and the smoother fields in ERA5-Land. In MAM, negative differences dominate much of southern Canada and large parts of the continental United States, indicating generally wetter conditions in ERA5-Land during spring, whereas positive differences persist over high latitudes. In JJA, negative differences extend over most of Canada, suggesting systematically larger summer accumulations in ERA5-Land, while CaPA-24h remains wetter over Mexico and parts of Central America. Persistent positive anomalies over Mexico throughout the year point to limitations near the southern edge of the domain.

The bottom two rows of Figure 7 compare CaPA-24h v3.2 and ERA5-Land against PRISM in terms of relative seasonal precipitation differences. CaPA-24h v3.2 generally shows closer agreement with PRISM than ERA5-Land, reflecting both the assimilation of similar station networks in CaPA-24h and PRISM and CaPA-24h enhanced ability to capture finer-scale spatial variability relative to the smoother ERA5-Land fields. However, both products exhibit pronounced discrepancies over the western United States, particularly along the Pacific coast and across major mountain ranges, reflecting the strong sensitivity of relative errors to complex topography. In these regions, CaPA-24h v3.2 generally shows positive differences relative to PRISM over windward slopes, while localized negative differences appear, especially in winter and spring. Large relative differences are also apparent over parts of the central United States, notably during winter and spring. These signals often coincide with regions of low seasonal precipitation totals (typically below 50 mm; first row in Fig. 7), such that modest absolute differences can translate into large relative values. Consequently, relative differences in these areas should be interpreted with caution; complementary maps of absolute differences are provided in the Supplemental Material (Fig. S2). Overall, CaPA-24h v3.2 exhibits finer spatial variability and more localized structures than ERA5-Land when compared to PRISM, particularly over complex terrain. While neither product consistently outperforms the other everywhere, CaPA-24h v3.2 tends to reproduce sharper gradients and regional contrasts that are more consistent with the high-resolution PRISM fields, whereas ERA5-Land shows smoother and more spatially homogeneous patterns across seasons. In summer (JJA), both CaPA-24h v3.2 and ERA5-Land show a systematic underestimation of seasonal precipitation accumulations relative to PRISM over large portions of the CONUS. This signal persists despite the higher density of precipitation observations assimilated in CaPA-24h during the warm season, indicating that it is unlikely to be primarily driven by observational coverage. Instead, it likely reflects limitations common to both reanalyses, including the representation of warm-season convective precipitation, scale mismatches between grid-cell averages and the high-resolution PRISM estimates, and smoothing inherent to the analysis systems. The underestimation is most apparent in regions where convective precipitation can dominate summer totals.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

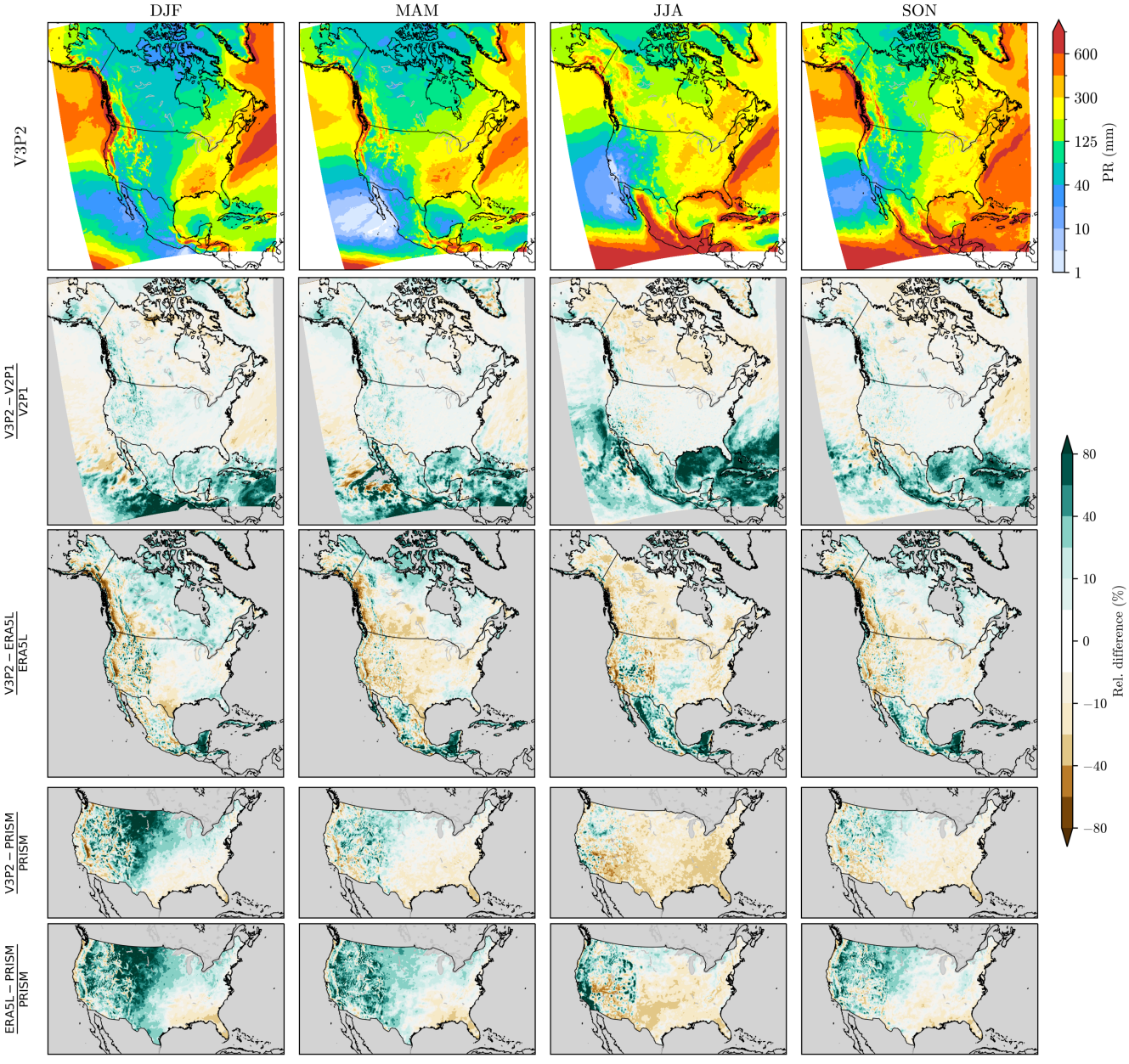

Figure 7. Seasonal mean precipitation accumulation and relative differences over 2002–2018. Top row: Seasonal precipitation accumulation (mm) from CaPA-24 v3.2. Remaining rows: Relative differences (%) in seasonal accumulation between (from top to bottom): CaPA-24h v3.2 and v2.1, CaPA-24h v3.2 and ERA5-Land, CaPA-24h v3.2 and PRISM, and ERA5-Land and PRISM. Each column corresponds to a season: winter (DJF), spring (MAM), summer (JJA), and autumn (SON). Relative differences are expressed as a percentage of the reference dataset in each row and are provided in the y-label.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

To further investigate the origin of these seasonal accumulation biases, the total precipitation bias relative to PRISM is decomposed into contributions from the frequency of wet days (precipitation > 1 mm) and the mean intensity on wet days. This decomposition also reveals regions where apparently small total biases result from compensating errors between precipitation frequency and intensity.

Figure 8 presents the regional decomposition of the total seasonal precipitation bias, computed with respect to PRISM, into contributions from wet-day frequency and precipitation intensity on wet days for CaPA-24h v3.2 and ERA5-Land across selected Bukovsky regions. Across most regions and seasons, differences in total accumulation are primarily driven by biases in wet-day frequency rather than by systematic errors in intensity. In winter, ERA5-Land exhibits a pronounced positive frequency bias over the Mountain West, the Great Plains, and Desert regions, which largely explains its positive total bias in these areas, while intensity biases remain comparatively modest. In contrast, CaPA-24h v3.2 shows smaller and more balanced frequency biases across regions, with intensity biases generally close to zero, indicating a better balance between occurrence and intensity errors.

In summer, it is noteworthy that in the South and Great Lakes regions the bias decomposition highlights a marked compensation between frequency and intensity errors in ERA5-Land. In these regions, ERA5-Land tends to overestimate the frequency of wet days while simultaneously underestimating precipitation intensity, resulting in near-zero total seasonal biases when only accumulated precipitation is considered. This compensation can give the impression of overall agreement with PRISM when examining total accumulations alone, despite substantial and physically meaningful discrepancies in the underlying precipitation characteristics. By contrast, CaPA-24h v3.2 exhibits smaller and more coherent biases in both frequency and intensity over the South and Great Lakes, leading to total biases that more directly reflect the underlying error structure. For the other regions, similarly to winter, CaPA-24h v3.2 shows a more balanced frequency–intensity bias compared to ERA5-Land, particularly over the Pacific region.

6.4 Agreement and discrepancies in extreme precipitation indices

KGE maps and their components for the annual daily maximum precipitation (Rx1day) are presented in Figure 9 for each season. The analysis focuses on the CONUS domain, where PRISM is available as reference, and on the 2002–2018 period.

A clear east–west contrast emerges in the KGE fields. East of the major mountainous regions, CaPA-24h v3.2 generally outperforms ERA5-Land in all seasons, with KGE values mostly ranging between 0.7 and 0.9. These high scores are mainly driven by bias and correlation components close to one, while the variability term is noisier, showing several blue patches indicative of higher variability in CaPA-24h relative to PRISM. In contrast, ERA5-Land systematically yields lower KGE values in this region, driven by larger biases (consistent with the accumulation results), weaker correlations, and reduced variability compared to PRISM. During summer, both reanalyses perform less well, with noisier KGE patterns, consistent with the challenges of capturing sub-grid convective storms (Ebert et al., 2007). Nonetheless, CaPA-24h v3.2 maintains smaller biases and higher correlations than ERA5-Land, which even exhibits locally negative correlations. Variability remains particularly noisy in both datasets but tends to be systematically underestimated in ERA5-Land.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

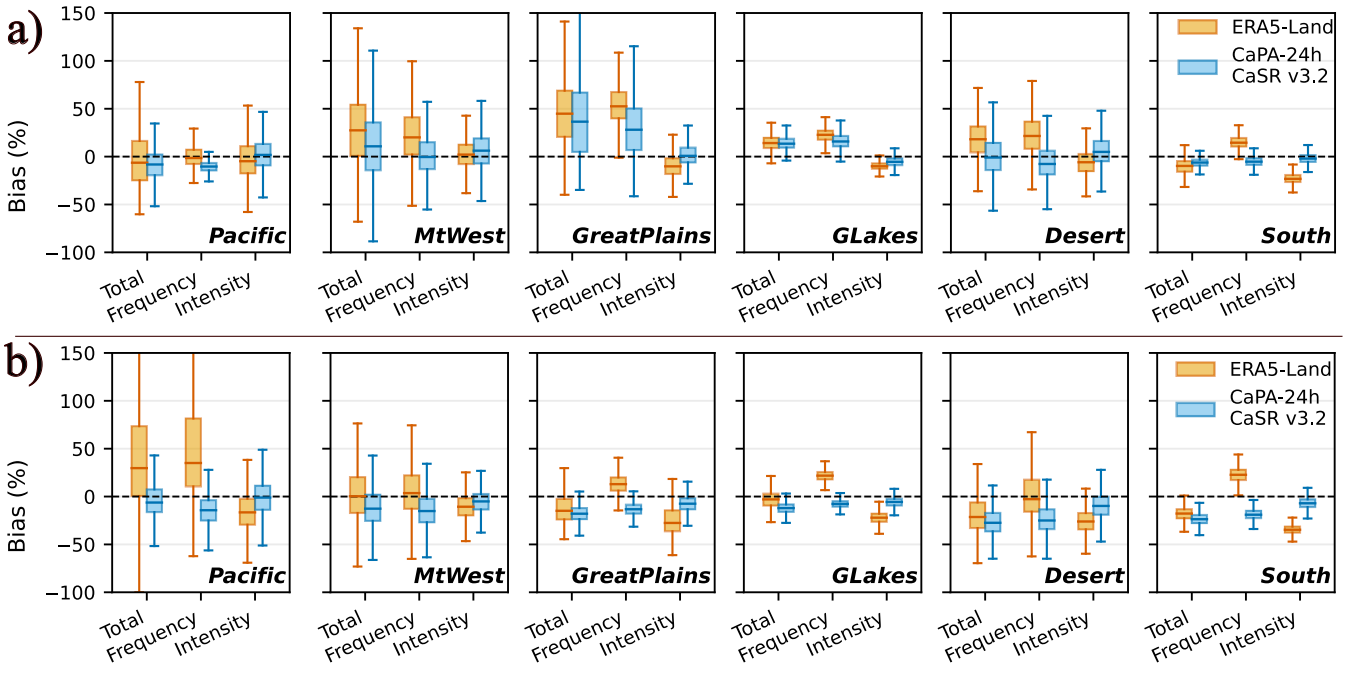

Figure 8. Decomposition of the seasonal precipitation bias relative to PRISM into contributions from wet-day frequency (precipitation > 1 mm) and mean precipitation intensity on wet days for ERA5-Land and CaPA-24h v3.2 over selected Bukovsky regions. Boxplots show the distribution of regional biases for (a) winter (DJF) and (b) summer (JJA) over the 2002–2018 period. Positive values indicate an overestimation by ERA5-Land or CaPA-24h v3.2 relative to PRISM.

In the western mountainous regions, performance degrades for both reanalyses, particularly with respect to bias. Across all seasons, CaPA-24h v3.2 exhibits patches of very low KGE values (below -0.4), especially in winter and spring. Following the interpretation of Knoben et al. (2019), such values indicate that the reanalysis does not add skill compared to the climatological mean. In these regions, both bias and variability contribute to the degradation of KGE.

ERA5-Land fields appear markedly smoother than those of CaPA-24h v3.2, consistent with the coarser-scale nature of ERA5, from which ERA5-Land directly inherits its characteristics without additional downscaling or bias correction (Muñoz-Sabater et al., 2021).

The same analysis applied to R95pTOT (Figure 10) complements Rx1day and other extreme precipitation metrics by assessing whether annual precipitation totals are dominated by very wet days or by light-to-moderate precipitation. KGE values for R95pTOT are generally lower than for Rx1day in both reanalyses. CaPA-24h v3.2 again outperforms ERA5-Land, particularly in MAM, JJA, and SON, while DJF remains the most challenging season to represent accurately.

Regional differences are evident. Both reanalyses perform poorly in the western mountainous regions, and ERA5-Land shows additional deficiencies over the eastern Appalachians, with KGE values below -0.4. In contrast, the southern part of the domain is relatively well represented, with KGE values around 0.7 or higher. Summer and autumn are the best-represented sea-

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

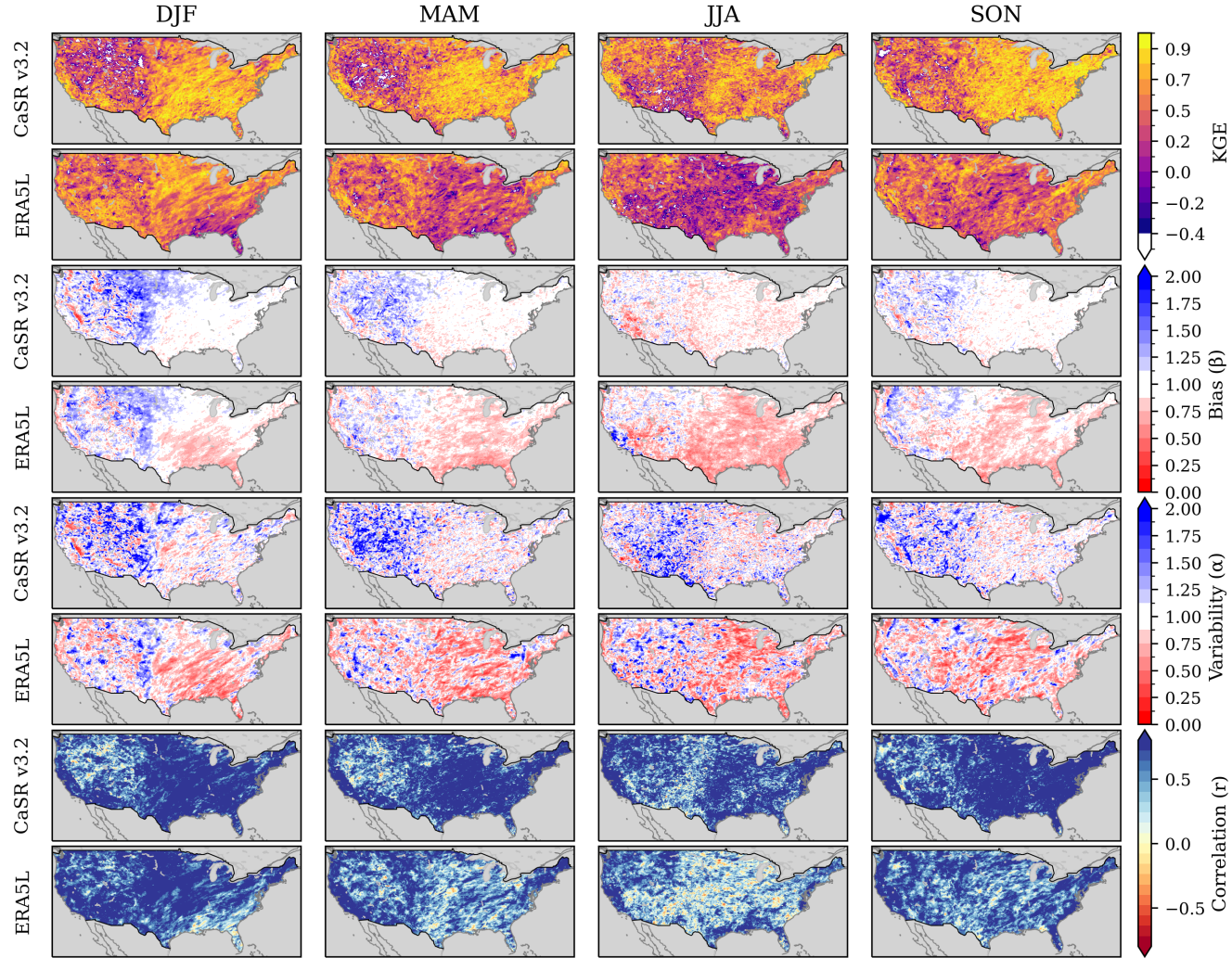

Figure 9. Kling–Gupta Efficiency (KGE) and its three components – bias ratio (β), variability ratio (α), and correlation (r) – for the annual **Rx1day** index, computed separately for each season (columns) and for the 2002-2018 period. Every other row corresponds to a different reanalysis, with CaPA-24h v3.2 and ERA5-Land evaluated against PRISM. The y-axis labels indicate the reanalysis product being compared. For all four metrics, a value of one represents a perfect match with PRISM.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

sons for CaPA-24h v3.2, coinciding with more active precipitation regimes (summer convection and autumn cyclones), where the assimilation of surface data likely provides added value, especially in the eastern domain, where the surface observation network is denser.

Low KGE values are primarily driven by large biases, with both CaPA-24h v3.2 and ERA5-Land overestimating mean R95pTOT by at least a factor of two in several regions and seasons. This result is somewhat counterintuitive, as one might expect mean annual R95pTOT values to be lower in the reanalyses given the finer native grid resolution of PRISM. However, because R95pTOT is defined relative to the 95th percentile, the higher threshold in PRISM leads to fewer exceedances and thus lower values. Conversely, the lower 95th-percentile threshold in CaPA-24h v3.2, and especially in the smoother ERA5-Land, produces more frequent exceedances. It should also be noted that percentiles are computed over the relatively short 2002–2018 period, which may limit their robustness. On the positive side, correlations for each season are close to one, indicating that the temporal variability of R95pTOT is well captured, particularly by CaPA-24h v3.2 and, to a lesser extent, ERA5-Land. This suggests that, with appropriate bias correction, the temporal evolution of R95pTOT could be accurately represented.

KGE and its components were also computed for Rx5day and R99pTOT (not shown; see Supplement). Regional and seasonal patterns are consistent with those obtained for Rx1day and R95pTOT, although performance generally decreases for these more extreme indices, which remain more challenging to reproduce in reanalyses.

6.5 Representation of the daily cycle of precipitation intensity

Analyzing the diurnal cycle of hourly precipitation provides insights beyond mean totals, as it reveals whether models capture the correct timing, amplitude, and regional structure of precipitation events. This diagnostic is particularly relevant in summer, when convection dominates (Pradhan et al., 2025; Dai, 2024), as many models tend to trigger rainfall too early in the day or misrepresent its intensity. In the following, the diurnal cycle is evaluated by comparing the hourly reanalysis from CaSR v3.2 (i.e. the disaggregated CaPA-24h, section 4) with ERA5-Land.

Figure 11 shows the median and interquartile range of the diurnal cycle of precipitation intensity for selected regions, displayed from 12 UTC to 12 UTC over the 1980–2018 period, for winter (Fig. 11a) and summer (Fig. 11b). In winter, as expected, the diurnal cycle exhibits weak variability across most regions, except in the Great Lakes region, where precipitation activity tends to peak during the morning local time in both CaSR v3.2 and ERA5-Land, and in the East region (not shown), where activity mainly occurs overnight and in the early morning, with a relative minimum around midday. Precipitation hourly rates are of similar magnitude in both datasets. These findings are consistent with Dai et al. (1999), who reported weak wintertime diurnal variations over the contiguous United States. Small peaks appear in CaSR v3.2 at specific synoptic hours (06, 12, and 18 UTC), corresponding to the lead-time stitching inherent to the disaggregation process. These artefacts are more pronounced in wetter regions, such as the South and, to a lesser extent, the Desert.

During summer (Fig. 11b), the diurnal cycle is more pronounced, and the non-physical discontinuities in CaSR v3.2 are more visible. Ignoring these artificial peaks, both datasets reproduce the expected pattern: summer precipitation generally peaks in the late afternoon local time (22–00 UTC) over the EBoreal, MtWest, Desert, Great Lakes, South, and East (not shown) regions. Some areas, such as the Prairies, display a secondary maximum around midnight (06 UTC) in addition to a main late-afternoon

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

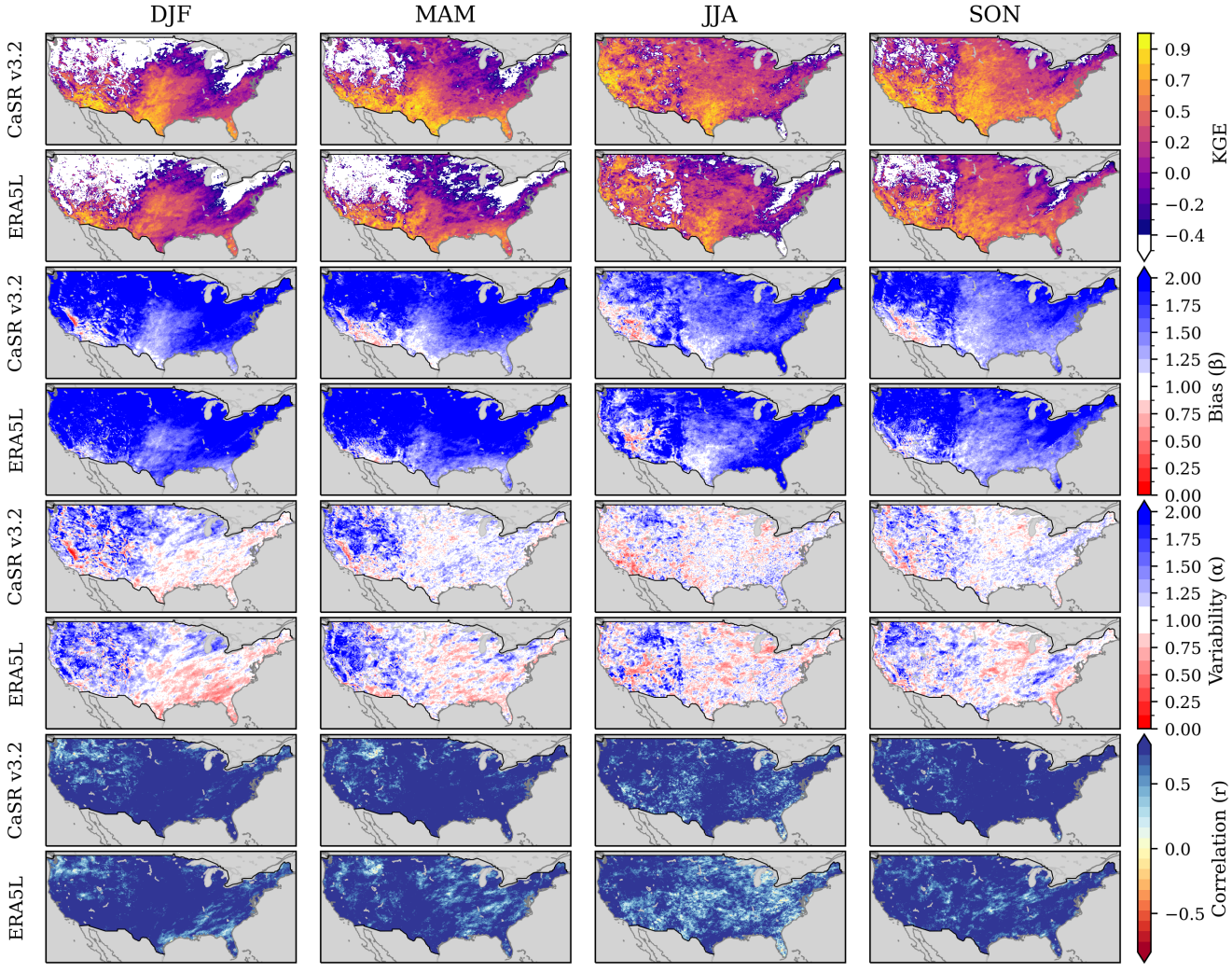

Figure 10. Same as Fig.9 but for the **R95pTOT** index.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

peak (23 UTC). These patterns are consistent with the diurnal cycles documented by Dai et al. (1999), highlighting the typical late-afternoon maximum of convective precipitation.

The artificial discontinuities in CaSR v3.2 at synoptic hours reach amplitudes of about 0.02 mm h^{-1} in summer and are associated with decreases in precipitation intensity immediately following each peak. These features partly reflect the lead-time stitching inherent to the temporal disaggregation procedure, whereby precipitation amounts tend to be enhanced near forecast boundaries. In addition, part of the bias in the diurnal cycle may be related to model spin-up effects, as the atmospheric model typically requires 12 to 24 hours to establish a dynamically consistent diurnal cycle of precipitation over the domain. During this adjustment period, precipitation characteristics have not yet fully stabilized and may differ from their representative diurnal cycle. The presence of these artefacts, and potential approaches to mitigate them, are discussed in the Conclusion as part of future improvements to the reanalysis. Overall, the temporal disaggregation procedure (Section 4) does not appear to substantially alter the phase of the diurnal cycle itself, although event-based analyses (Dunkerley, 2008) could provide further insight.

7 Near-real-time completion of CaPA-24h from CaSR v3.2 using operational CaPA

As described in Sect. 4, the CaSR v3.2 reanalysis currently extends until December 2024. For users seeking to extend the precipitation time series beyond this period using the operational CaPA system, it is important to understand the main similarities and differences between the two products, as well as the precautions required when combining them. The recommended methodology for constructing CaSR-like precipitation fields from operational CaPA for near–real-time applications is documented online (https://hpfx.collab.science.gc.ca/~scar700/rcas-casr/dataset_specifics.html#create_casr-like_with_ops).

Among the different operational versions of CaPA (low-resolution, high-resolution, and ensemble), the 24-hour accumulation from the Regional Deterministic Precipitation Analysis (hereafter CaPA-RDPA) is considered, as it most closely resembles CaPA-24h in terms of grid spacing ($\Delta \simeq 10$ km) and domain coverage (North America). Background field of this version of CaPA is based on an identical model configuration but differ by the initialization and lateral boundaries. Both the operational and CaSR v3.2 grids are geographically collocated, although the domain limits differ slightly. The current operational configuration of CaPA-RDPA is documented in Lespinas et al. (2024), and its main characteristics compared to CaPA-24h (CaSR v3.2) are summarized in Table 1.

The shorter background forecast window used in CaPA-RDPA (06–12 UTC) compared to CaPA-24h (06–18 UTC), together with the assimilation of radar-based QPE and IMERG precipitation estimates, is expected to introduce systematic differences between the two analyses. To assess the impact of these differences, the operational CaPA-RDPA configuration was compared with CaSR CaPA-24h using standard categorical verification metrics (i.e., FBI and ETS), as well as seasonal accumulation maps and time series. The evaluation was conducted over the 2021–2022 water year (October 2021–September 2022), corresponding to the most recent operational configuration available as of June 2024. FBI and ETS were computed using manual synoptic stations in winter and a combination of manual and automatic stations in summer, with an identical verification station sample applied to both experiments.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

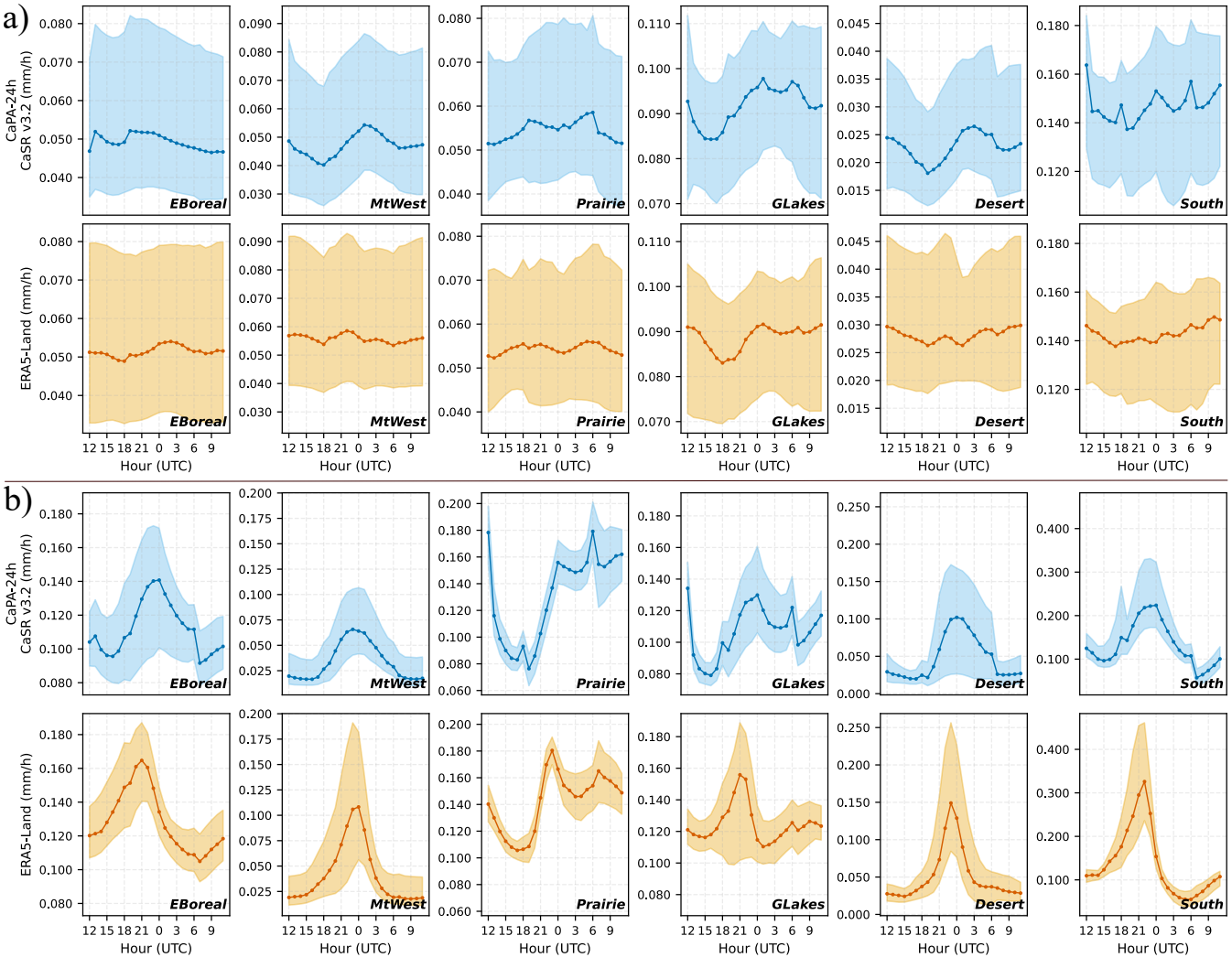

Figure 11. Diurnal cycle of hourly precipitation intensity in winter (a, DJF) and summer (b, JJA) for CaSR v3.2 (blue) and ERA5-Land (gold). The thick line shows the regional median of the hourly mean precipitation over 1980–2018, and the shaded area denotes the interquartile range (25th–75th percentile).

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

Table 1. Main characteristics of the 24-hour CaPA-RDPA and CaPA-24h (CaSR v3.2) analyses. RDOS refers to the Regional Deterministic Prediction System, a numerical weather prediction model operated at ECCC. Radar QPE and IMERG are assimilated only during liquid-precipitation events (temperature threshold of 0 °C).

	CaPA-RDPA	**CaPA-24h (CaSR v3.2)**
Spatial resolution	∼10 km	∼10 km
Domain/ grid size	North America 1140 × 1045	North America 706 × 778
Validity time	06 and 12 UTC	12 UTC
Assimilated inputs	Surface stations Radar QPE IMERG	Surface stations
Lead time window	06–12 UTC	06–18 UTC
Background NWP model	RDPS (based on GEM 5.2.1) issued at 0, 6, 12 & 18 UTC	RDRS (based on GEM 5.2.1) issued at 0 & 12 UTC
Global initial conditions	GDPS (GEM 5.2.1)	ERA5
System type	Operational (updated every 3 to 4 years)	Static reanalysis version

Figure 12 compares the performance of CaPA-24h (CaSR v3.2) and the operational CaPA-RDPA in terms of FBI and ETS. In winter, when neither radar nor IMERG data are assimilated, both systems exhibit approximately 30 % more precipitation events than observed, as indicated by the FBI. This overestimation decreases to roughly 20 % in summer. For light precipitation (below 5 mm d^{-1}), both analyses reproduce the observed frequency reasonably well, whereas at higher thresholds they tend to underestimate the number of events, particularly in summer, with approximately 25 % (CaPA-24h) to 50 % (RDPA) fewer occurrences than observed. The background fields exhibit larger frequency biases in CaSR v3.2 than in RDPA. This difference is consistent with the longer 06–18 UTC background forecast window used in the reanalysis, which includes a longer model adjustment (spin-up) phase at the beginning of each cycle. During this period, precipitation characteristics are still stabilizing, which can contribute to frequency biases in the background field. In both systems, the analysis step substantially reduces these biases in all seasons, although residual deficiencies remain for high-intensity precipitation events.

In terms of ETS, both analyses perform similarly in winter, maintaining scores around 0.6 for the 5–20 mm threshold range before declining rapidly at higher thresholds. In summer, RDPA achieves slightly higher scores despite comparable background quality, though both systems show reduced performance relative to winter. An additional RDPA experiment assimilating surface stations only, without radar or IMERG inputs (not shown), produced nearly identical ETS scores to those of CaPA-24h, demonstrating that radar QPE and IMERG primarily enhance the spatial and temporal coherence of the analyzed precipitation fields. However, the FBI-1 remained nearly unchanged, suggesting that these additional datasets mainly improve where precipitation occurs rather than how frequently.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

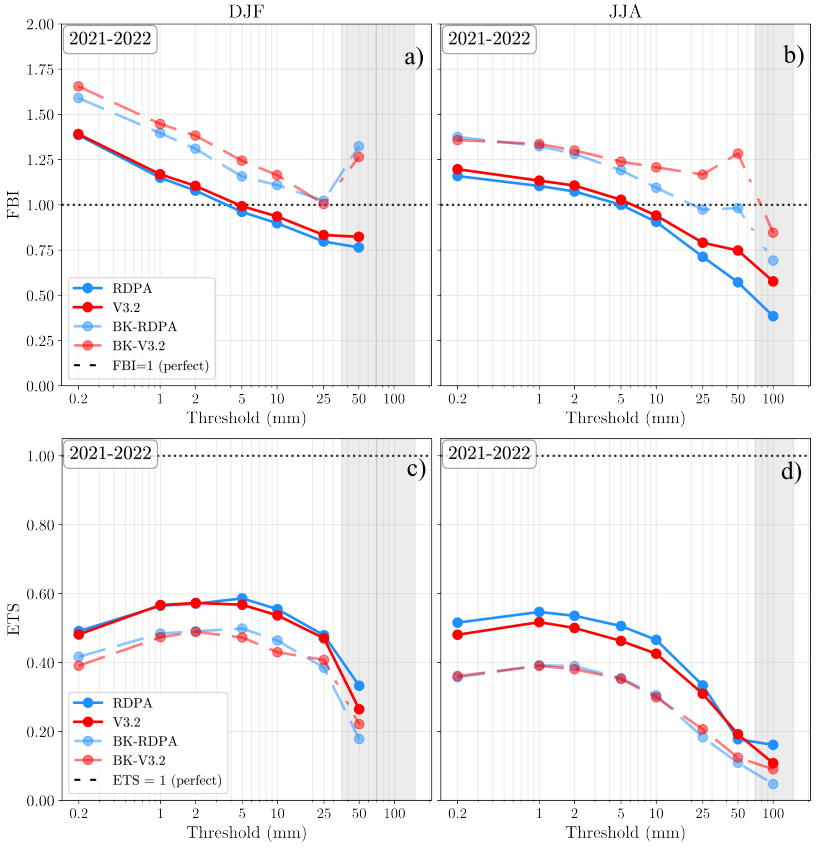

Figure 12. Performance of daily precipitation estimates in terms of the Equitable Threat Score (ETS) and Frequency Bias Index (FBI) for winter (DJF; panels a and c) and summer (JJA; panels b and d). Results are shown for CaPA-24h v3.2 (red circles), CaPA-RDPA (blue circles), and their respective background fields (BK–v3.2 and BK–RDPA; dashed lines in matching colors). The shaded gray area indicates thresholds where the sample size is too small to ensure meaningful statistics (< 50 of hits and misses).

Figure 13 shows the seasonal accumulated precipitation for winter (DJF) and summer (JJA) of 2022 from both CaPA-24h (v3.2) and the operational RDPA. In winter, both analyses display very similar large-scale patterns, indicating overall consistency between the two products. Some localized differences are visible, notably along the western cordillera, where CaPA-24h fields appear slightly smoother. In summer, larger differences emerge over eastern Canada, particularly in Québec, where RDPA shows lower accumulated precipitation than CaPA-24h. This reduction is linked to the assimilation of radar and IMERG data in the operational system, which tend to constrain convective precipitation signals. Indeed, when examining the equivalent RDPA experiment that assimilates only surface stations (not disseminated publicly), the differences over eastern Canada are substantially reduced.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

The time series shown in Figure 14 place these seasonal differences in a longer-term context and indicate that regional seasonal precipitation from CaPA-24h v3.2 and the operational RDPA remains highly consistent, with overlapping interquartile ranges and coherent interannual variability.

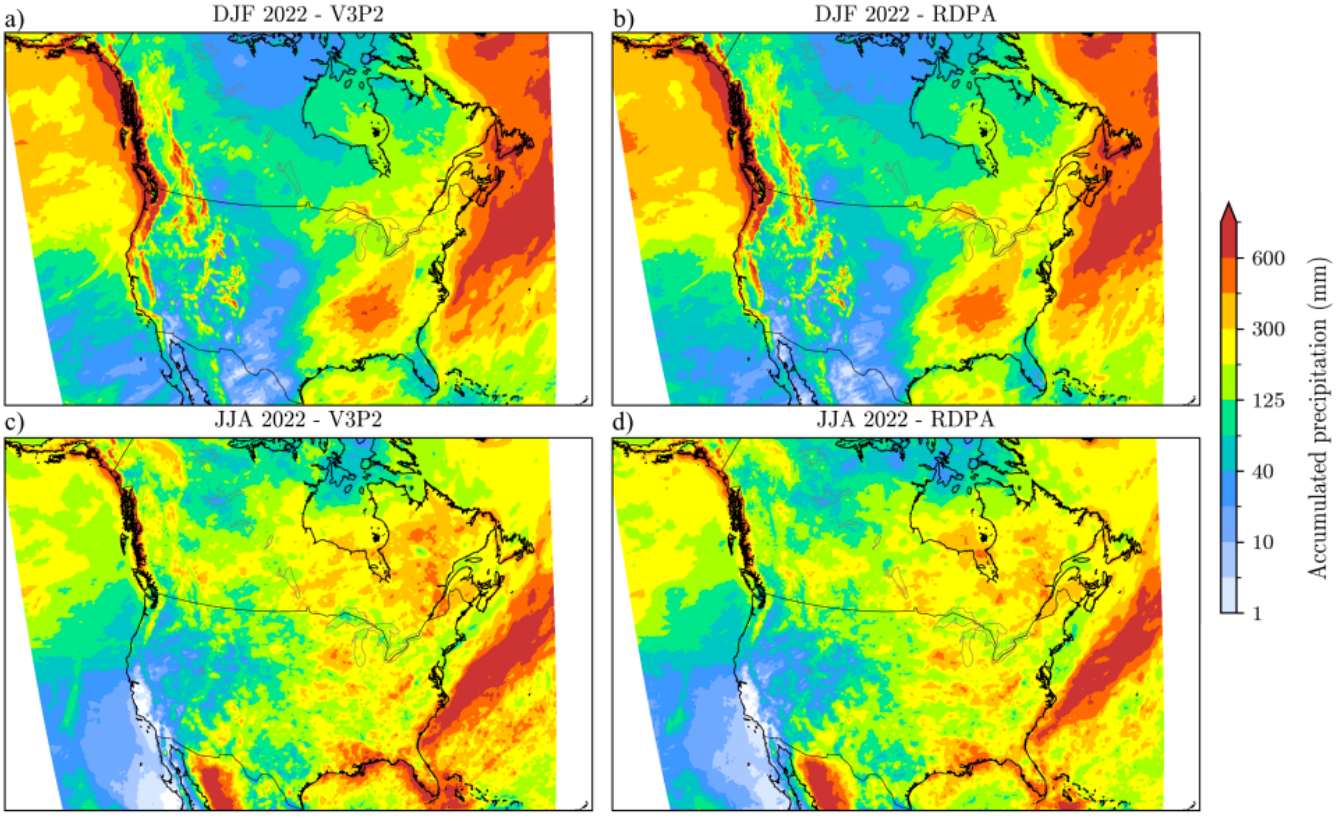

Figure 13. Seasonal accumulated precipitation for winter (DJF, panels a–b) and summer (JJA, panels c–d) of 2022 from (a, c) CaPA-24h within CaSR v3.2 and (b, d) the operational RDPA.

Taken together, these results indicate that, despite localized seasonal differences, the two systems are broadly consistent at climatological scales. Although this comparison is based on a single year of overlap, the results suggest that using operational CaPA-RDPA to extend CaSR v3.2 in near-real-time applications is appropriate for climatological analyses, such as regional trend assessments and seasonal to interannual variability studies, without introducing major discontinuities in long-term time series. However, caution is warranted when using operational RDPA for weather-scale applications or detailed event-based analyses, particularly during summer months and over regions where radar and satellite data are assimilated. In such cases, localized differences in precipitation amounts and the spatial distribution of convective events may be significant, and users should account for these systematic differences when interpreting results at finer temporal and spatial scales.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

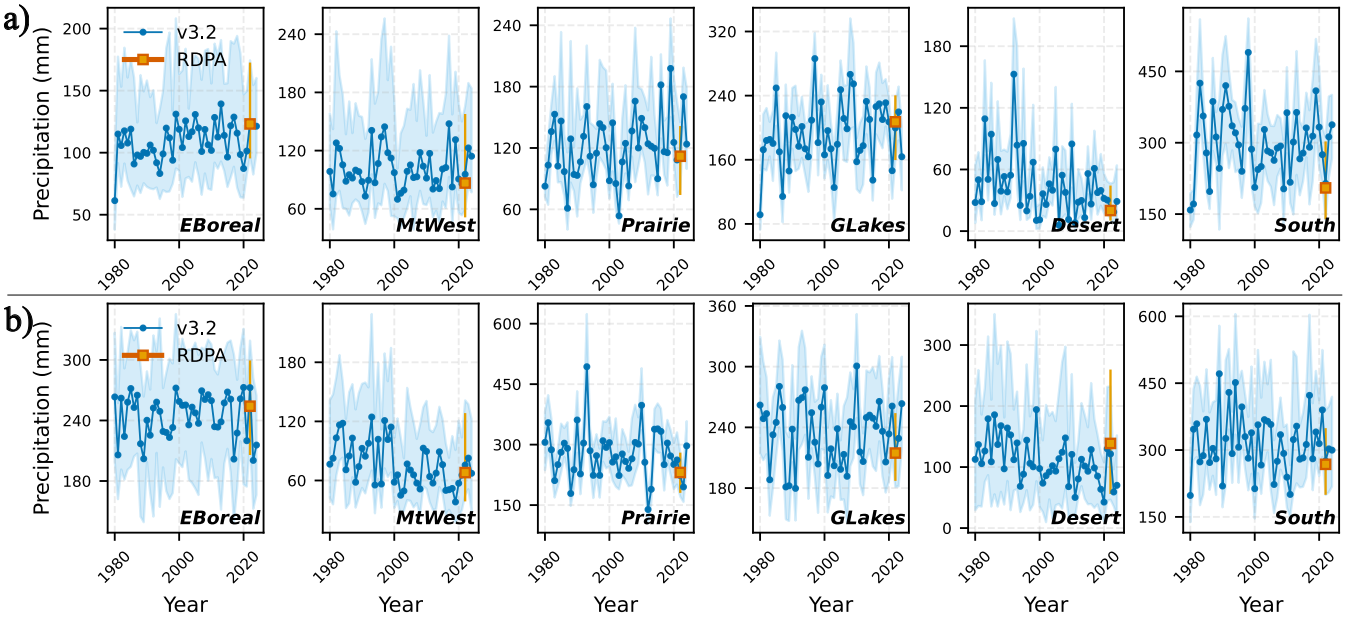

Figure 14. Time series of seasonal accumulated precipitation for selected regions in (a) winter (DJF) and (b) summer (JJA). The solid blue line and red marker show the regional medians for CaPA-24h from CaSR v3.2 and the operational RDPA, respectively. Shaded areas represent the interquartile range (25th–75th percentiles) across all grid points within each region.

8 Conclusions and discussions

CaSR v3.2 provides a high-resolution precipitation reanalysis for North America spanning 1980–2024, built upon a two-component system. The *online* component produces short-range forecasts from a regional numerical weather prediction configuration coupled to land-surface assimilation, while the *offline* component (CaPA-24h) generates an observation-driven 24-hour precipitation analysis at 12 UTC by optimally combining surface gauge observations with a background field derived from the online forecasts. In version 3.2, substantial upgrades to the GEM model physics and resolution, together with improved observation handling and a denser, better curated gauge archive, translate into more reliable precipitation background fields and, consequently, improved offline analyses.

This study evaluated CaPA-24h v3.2 relative to its predecessor (v2.1) and to two independent gridded datasets, ERA5-Land (Muñoz-Sabater et al., 2021) and PRISM (Daly et al., 2008). Verification against in situ observations relied on leave-one-out analyses embedded within CaPA to limit artificial skill inflation, and used complementary categorical and intensity diagnostics (FBI, ETS, and the partial mean). Intercomparisons of gridded fields further assessed seasonal accumulations, a frequency–intensity decomposition of total biases relative to PRISM, and extreme precipitation indices (Rx1day and R95pTOT) evaluated using the Kling–Gupta Efficiency (KGE). In addition, the hourly precipitation product derived from temporal disaggregation of CaPA-24h was examined through regional diagnostics of the diurnal cycle. The main conclusions are as follows:

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

Several consistent features emerge from these analyses. Station-based verification shows that CaPA-24h v3.2 achieves a level of skill comparable to that of v2.1 at observation locations, reflecting the strong constraint imposed by assimilated gauges. At the same time, the background field in v3.2 exhibits systematically reduced frequency bias and improved intensity characteristics relative to v2.1, differences that are most influential in data-sparse regions where the background field contributes more strongly to the final analysis. When the evaluation is separated into the 1980–1999 and 2000–2018 periods, the relative ranking of the products remains stable, and only limited changes in performance are observed. Differences between the two periods are mainly confined to moderately extreme thresholds and are consistent with sampling effects related to lower station density prior to 2000 and reduced wintertime gauge availability, rather than with fundamental changes in system behaviour.

Gridded intercomparisons over the 2002–2018 period show that CaPA-24h v3.2 shows closer agreement with PRISM than ERA5-Land for seasonal accumulation patterns across much of the eastern domain, while both products exhibit their largest discrepancies over complex terrain. The frequency–intensity decomposition demonstrates that biases in seasonal totals are most often driven by errors in wet-day frequency rather than by systematic intensity errors. This analysis also highlights cases in which ERA5-Land exhibits compensating frequency and intensity biases, which can mask deficiencies when only total accumulations are considered. Finally, extreme precipitation indices display a pronounced east–west contrast: CaPA-24h v3.2 generally achieves higher KGE values than ERA5-Land over the eastern CONUS, whereas performance degrades for both products over the western cordillera, where remaining orographic biases strongly affect both mean and variability.

The hourly precipitation fields derived from temporal disaggregation capture the expected seasonal contrast in diurnal variability and reproduce the late-afternoon maximum of warm-season precipitation over several regions. However, non-physical peaks at synoptic hours, followed by intensity drops, remain visible and reflect the current lead-time stitching and rescaling strategy. These artefacts are most pronounced in summer and in wetter regions, and may affect event-duration statistics and derived hydrological indicators.

Several limitations condition the interpretation and use of the dataset. Point–grid representativeness errors affect threshold-based verification, particularly for localized convective extremes and in regions of complex topography. Although the leave-one-out framework reduces the direct impact of assimilated observations on verification, spatial dependence among validation points cannot be fully eliminated. The pre-2000 evaluation period is also more sensitive to sampling, owing to reduced station density and more limited wintertime gauge availability. In addition, PRISM exhibits documented temporal inhomogeneities linked to changes in inputs and processing, motivating restriction of PRISM-based analyses to the 2002–2018 period and limiting the robustness of percentile-based extreme indices.

From a user perspective, CaPA-24h v3.2 can be used for climatological analyses over North America at seasonal to interannual time scales, as well as for regional water-cycle assessments and land-surface or hydrological modelling applications. Relative to ERA5-Land, differences in the representation of extreme precipitation are apparent, with spatial patterns indicating higher agreement with observational references over large parts of the domain and reduced performance in regions of complex topography. At sub-daily resolution, hourly precipitation fields reflect limitations inherent to the current disaggregation methodology, which should be considered when analysing the diurnal cycle or event-scale precipitation statistics.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

To ensure continuity beyond December 2024, CaPA-24h from CaSR v3.2 was compared with the operational CaPA-RDPA over the 2021–2022 water year. Despite localized seasonal differences—most notably in summer over regions influenced by radar and satellite assimilation in the operational system—the two products remain broadly consistent at climatological scales. This consistency supports the use of operational CaPA-RDPA to extend CaSR v3.2 time series for climate-oriented applications, while continued caution is advised for weather-scale or event-based analyses, particularly during warm-season convective regimes.

Future developments should prioritize directions that address the remaining limitations identified here. These include reducing spin-up effects through more frequent model initialization, improving daily-to-hourly disaggregation strategies to mitigate synoptic-hour artefacts, integrating high-resolution radar and satellite precipitation products, and exploring machine-learning-based bias correction approaches. Together, these advances are expected to further enhance the reliability of CaSR precipitation fields, particularly for extremes and sub-daily applications.

Data availability. The CaSR v3.2 datasets are publicly available through the Environment and Climate Change Canada data portal at https://hpfx.collab.science.gc.ca/~scar700/rcas-casr/. In addition to the core CaSR atmospheric and precipitation products, two derived datasets are distributed: CaSR-Land, providing a surface reanalysis, and CaSR-River, providing a river discharge reanalysis based on CaSR forcings. An intermediate release (CaSR v3.1) was previously disseminated. This version contained a known issue affecting the offline precipitation reanalysis over the province of Québec. The issue has been fully corrected in CaSR v3.2, which supersedes all earlier v3 releases and should be used for scientific applications.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

Appendix A: Background field and details of input datasets

A1 Construction of the 24-hour background field

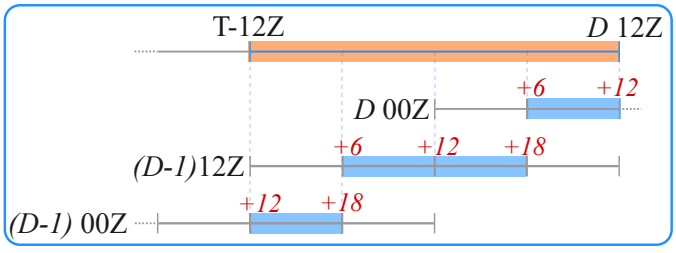

Figure A1. Selected lead times from RDRS (in shaded blue) 24-hour issued forecast to build the background field for CaPA-24h valid at 12Z on day D (in shaded orange).

A2 Details of assimilated datasets

Table A1. Observational networks assimilated in CaPA reanalysis, along with their periods of availability. Variables used include precipitation (P) but also temperature (T) and wind speed ($||U||$). T and $||U||$ are not assimilated but are used for quality control purposes. Data origin is indicated in parentheses.

Network	Domain	Period (Origin)	CaPA-24h
SYNOP	North America	1980–1999 (ISD)	$P, T, ||U||$
		2000–present (ECCC)	–
SWOB	North America	1980–1999 (ISD)	$P, T, ||U||$
		2013–present (ECCC)	–
METAR	North America	1980–1999 (ISD)	$P, T, ||U||$
		2000–present (ECCC)	–
AdjDlyRS	Canada	1980–2023 (ECCC)	P
AdjHlyRS	Canada	2001–2023 (ECCC)	P
RMCQ	Quebec	2011–present (ECCC)	$P, T, ||U||$
SHEF	USA	2000–present (ECCC)	$P, T, ||U||$
CoCoRaHS	Canada–USA	(ECCC)	P

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

Appendix B: Regions over the North American domain

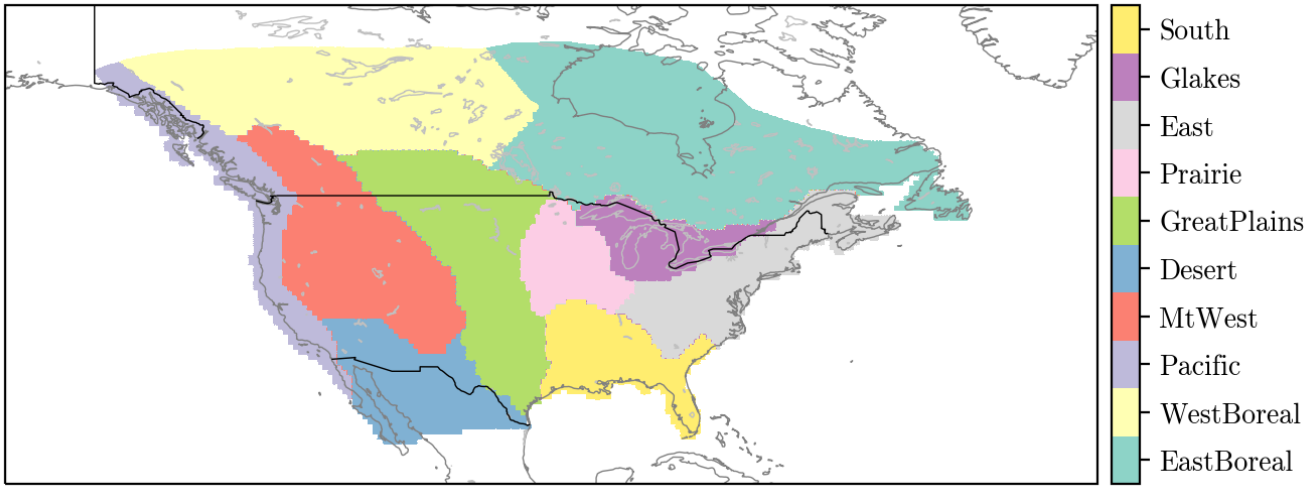

Figure B1. Regions used in this study following the division proposed by Bukovsky (2011)

Appendix C: Accumulations across the western mountainous region

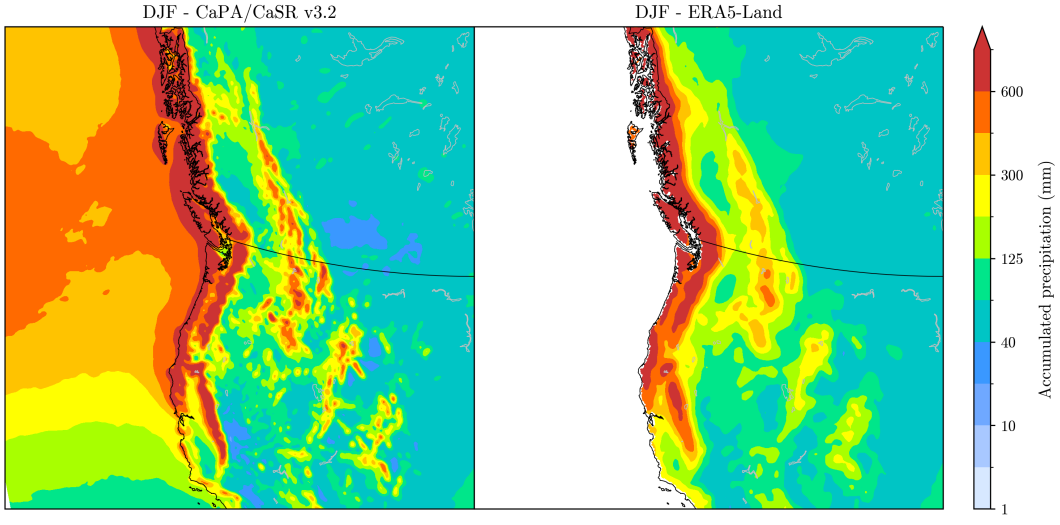

Figure C1. Mean winter accumulated precipitation (PR in mm) over the 1981–2018 period, shown for CaSR v3.2 CaPA-24h (left) and ERA5-Land (right).

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

Appendix D: Scores

The 2x2 contingency table for binary events for precipitation verification is shown in Table D1 below.

Table D1. The 2×2 contingency table for binary event verification. The elements a, b, c, and d represent the frequency of hits (correct detections), false alarms (incorrect detections), misses (undetected events), and correct negatives (correct rejections), respectively.

	Observed event Yes	Observed event No	Total
Analysis event Yes	a (hit)	b (false alarm)	a+b
Analysis event No	c (miss)	d (correct negative)	c+d
Total	a+c	b+d	1

The Frequency Bias Index (FBI) and Equitable Threat Scores (ETS) based on Table D1 expresses as:

$$\mathrm{FBI} = \frac{(a+b)}{(a+c)}, \tag{D1}$$

$$\mathrm{ETS} = \frac{a - h_r}{a+b+c-h_r}, \tag{D2}$$

where $h_r = (a+b)\,(a+c)$ illustrates the hits expected by chance.

The partial mean is given by the following:

$$\mathbb{E}[S \mid S < q] = \frac{\sum_{i=1}^{N} S_i \, \mathbf{1}_{\{S_i < q\}}}{\sum_{i=1}^{N} \mathbf{1}_{\{S_i < q\}}}, \tag{D3}$$

where S_i is the daily precipitation of a given dataset and $\mathbf{1}$ is the indicator function that maps daily precipitation to 1 if they are strictly below a given threshold q.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

Author contributions. DK wrote the article, generated the figures and did the analysis of the results. NG and DK participated to the elaboration of the running of CaPA system in the reanalysis context. NG was also responsible of preparing the set-up to properly run the CaSR online component over the 1980-2024 period. VF participated to the review of the article and discussed results. MD contributed to the evaluations and additional seasonal quality control checks of the versions v2 and 3 and the review of the article. MB was responsible for the preparation of the in situ datasets used of all components of CaSR. XW run the distributed CaPA-24h over the period.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Financial support from Natural Resources Canada for CaSR version 3 through the Flood Hazard Identification and Mapping Program (FHIMP) is gratefully acknowledged. The authors also acknowledge the International Joint Commission, through its International Watersheds Initiative, for its contribution to the initiation of the CaSR project leading to the current version.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

References

Baatz, R., Franssen, H. J. H., Euskirchen, E., Sihi, D., Dietze, M., Ciavatta, S., Fennel, K., Beck, H., Lannoy, G. D., Pauwels, V. R. N., Raiho, A., Montzka, C., Williams, M., Mishra, U., Poppe, C., Zacharias, S., Lausch, A., Samaniego, L., Looy, K. V., Bogena, H., Adamescu, M., Mirtl, M., Fox, A., Goergen, K., Naz, B. S., Zeng, Y., and Vereecken, H.: Reanalysis in Earth System Science: Toward Terrestrial Ecosystem Reanalysis, Reviews of Geophysics, 59, https://doi.org/10.1029/2020RG000715, publisher: John Wiley and Sons Inc, 2021.

Beck, H. E., Dijk, A. I. J. M. v., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and Roo, A. d.: MSWEP: 3-hourly 0.25 global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data, Hydrology and Earth System Sciences, 21, 589–615, https://doi.org/10.5194/hess-21-589-2017, 2017.

Bell, B., Hersbach, H., Simmons, A., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Soci, C., Villaume, S., Bidlot, J., Haimberger, L., Woollen, J., Buontempo, C., and Thépaut, J.: The ERA5 global reanalysis: Preliminary extension to 1950, Quarterly Journal of the Royal Meteorological Society, 147, 4186–4227, https://doi.org/10.1002/qj.4174, 2021.

Bengtsson, L. and Shukla, J.: Integration of Space and In Situ Observations to Study Global Climate Change, Bulletin of the American Meteorological Society, 69, 1130–1143, https://doi.org/https://doi.org/10.1175/1520-0477(1988)069<1130:IOSAIS>2.0.CO;2, 1988.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., and Tian, Q.: Accurate medium-range global weather forecasting with 3D neural networks, Nature, 619, 533–538, https://doi.org/10.1038/s41586-023-06185-3, 2023.

Bosilovich, M. G., Chen, J., Robertson, F. R., and Adler, R. F.: Evaluation of Global Precipitation in Reanalyses, Journal of Applied Meteorology and Climatology, 47, 2279–2299, https://doi.org/10.1175/2008JAMC1921.1, 2008.

Bukovsky, M. S.: Masks for the Bukovsky regionalization of North America., www.narccap.ucar.edu/contrib/bukovsky/, 2011.

Caron, J. F., Milewski, T., Buehner, M., Fillion, L., Reszka, M., Macpherson, S., and St-James, J.: Implementation of deterministic weather forecasting systems based on ensemble-variational data assimilation at Environment Canada. Part II: The regional system, Monthly Weather Review, 143, 2560–2580, https://doi.org/10.1175/MWR-D-14-00353.1, publisher: American Meteorological Society, 2015.

Carrera, M. L., Bélair, S., and Bilodeau, B.: The Canadian Land Data Assimilation System (CaLDAS): Description and Synthetic Evaluation Study, Journal of Hydrometeorology, 16, 1293–1314, https://doi.org/10.1175/JHM-D-14-0089.1, 2015.

Copernicus: Climate Data Store, https://cds.climate.copernicus.eu/.

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., and Buontempo, C.: WFDE5: bias-adjusted ERA5 reanalysis data for impact studies, Earth System Science Data, 12, 2097–2120, https://doi.org/10.5194/essd-12-2097-2020, 2020.

Dai, A.: The diurnal cycle from observations and ERA5 in precipitation, clouds, boundary layer height, buoyancy, and surface fluxes, Climate Dynamics, https://doi.org/10.1007/s00382-024-07182-6, 2024.

Dai, A., Giorgi, F., and Trenberth, K. E.: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States, Journal of Geophysical Research: Atmospheres, 104, 6377–6402, https://doi.org/10.1029/98JD02720, 1999.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., and Pasteris, P. P.: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, International Journal of Climatology, 28, 2031–2064, https://doi.org/10.1002/joc.1688, publisher: John Wiley and Sons Ltd, 2008.

Daly, C., Doggett, M. K., Smith, J. I., Olson, K. V., Halbeib, M. D., Dimcovic, Z., Keon, D., Loiselle, R. A., Steinbeirg, B., Ryan, A. D., Pancake, C. M., and Kaspar, E. M.: Challenges in Observation-Based Mapping of Daily Precipitation across the Conterminous United States, Journal of Atmospheric and Oceanic Technology, 38, 1979–1992, https://doi.org/10.1175/JTECH-D-21-0054.1, 2021.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J., Park, B., Peubey, C., De Rosnay, P., Tavolato, C., Thépaut, J., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Quarterly Journal of the Royal Meteorological Society, 131, 3385–3396, https://doi.org/10.1256/qj.05.108, 2005.

Devers, A., Vidal, J.-P., Lauvernet, C., and Vannier, O.: FYRE Climate: A high-resolution reanalysis of daily precipitation and temperature in France from 1871 to 2012, Climate of the Past, 17, 1857–1879, https://doi.org/10.5281/zenodo.4005573, 2021.

Dunkerley, D.: Identifying individual rain events from pluviograph records : a review with analysis of data from an Australian dryland site, Hydrological Processes, 22, 5024–5036, https://doi.org/10.1002/hyp, 2008.

Ebert, E. E., Janowiak, J. E., and Kidd, C.: Comparison of near-real-time precipitation estimates from satellite observations and numerical models, Bulletin of the American Meteorological Society, 88, 47–64, https://doi.org/10.1175/BAMS-88-1-47, 2007.

Fortin, V., Roy, G., Donaldson, N., and Mahidjiba, A.: Assimilation of radar quantitative precipitation estimations in the Canadian Precipitation Analysis (CaPA), Journal of Hydrology, https://doi.org/10.1016/j.jhydrol.2015.08.003, 2015.

Fortin, V., Roy, G., Stadnyk, T., Koenig, K., Gasset, N., and Mahidjiba, A.: Ten Years of Science Based on the Canadian Precipitation Analysis: A CaPA System Overview and Literature Review, Atmosphere-Ocean, 56, 178–196, https://doi.org/10.1080/07055900.2018.1474728, 2018.

Gasset, N. and Milewski, T.: Changes to the Regional Deterministic Prediction System (RDPS) from version 8.1.0 to version 9.0.0, Technical Note, ECCC, https://collaboration.cmc.ec.gc.ca/cmc/cmoi/product_guide/docs/tech_notes/technote_rdps_e.pdf, 2024.

Gasset, N., Fortin, V., Dimitrijevic, M., Carrera, M., Bilodeau, B., Muncaster, R., Gaborit, , Roy, G., Pentcheva, N., Bulat, M., Wang, X., Pavlovic, R., Lespinas, F., and Khedhaouiria, D.: A 10 km North American Precipitation and Land Surface Reanalysis Based on the GEM Atmospheric Model, Hydrology and Earth System Sciences, https://doi.org/10.5194/hess-2021-41, 2021.

Gehne, M., Hamill, T. M., Kiladis, G. N., and Trenberth, K. E.: Comparison of Global Precipitation Estimates across a Range of Temporal and Spatial Scales, Journal of Climate, 29, 7773–7795, https://doi.org/10.1175/JCLI-D-15-0618.s1, 2016.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Girard, C., Plante, A., Desgagné, M., McTaggart-Cowan, R., Côté, J., Charron, M., Gravel, S., Lee, V., Patoine, A., Qaddouri, A., Roch, M., Spacek, L., Tanguay, M., Vaillancourt, P. A., and Zadra, A.: Staggered Vertical Discretization of the Canadian Environmental Multiscale (GEM) Model Using a Coordinate of the Log-Hydrostatic-Pressure Type, Monthly Weather Review, 142, 1183–1196, https://doi.org/10.1175/MWR-D-13-00255.1, 2014.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, publisher: Elsevier B.V., 2009.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

Hamill, T. M. and Whitaker, J. S.: Probabilistic Quantitative Precipitation Forecasts Based on Reforecast Analogs: Theory and Application, Monthly Weather Review, 134, 3209–3229, https://doi.org/10.1175/MWR3237.1, 2006.

Hersbach, H.: Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems, Weather and Forecasting, 15, 559–570, https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2, 2000.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., De Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hu, G. and Franzke, C. L. E.: Evaluation of Daily Precipitation Extremes in Reanalysis and Gridded Observation-Based Data Sets Over Germany, Geophysical Research Letters, 47, e2020GL089 624, https://doi.org/10.1029/2020GL089624, 2020.

Jiawei Zhuang, raphael dussin, David Huard, Pascal Bourgault, Anderson Banihirwe, Stephane Raynaud, Brewster Malevich, Martin Schupfner, Filipe, Charles Gauthier, Sam Levang, André Jüling, Mattia Almansi, RichardScottOZ, RondeauG, Stephan Rasp, Trevor James Smith, Ben Mares, Jemma Stachelek, Matthew Plough, Pierre, Ray Bell, Romain Caneill, and Xianxiang Li: pangeo-data/xESMF: v0.8.10, https://doi.org/10.5281/ZENODO.4294774, 2025.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP–DOE AMIP-II Reanalysis (R-2), Bulletin of the American Meteorological Society, 83, 1631–1643, https://doi.org/10.1175/BAMS-83-11-1631, 2002.

Karl, T. R., Nicholls, N., and Ghazi, A.: CLIVAR/GCOS/WMO Workshop on Indices and Indicators for Climate Extremes Workshop Summary, in: Weather and Climate Extremes, edited by Karl, T. R., Nicholls, N., and Ghazi, A., pp. 3–7, Springer Netherlands, Dordrecht, ISBN 9789048152230 9789401592659, https://doi.org/10.1007/978-94-015-9265-9_2, 1999.

Khedhaouiria, D., Bélair, S., Fortin, V., Roy, G., and Lespinas, F.: Using a hybrid optimal interpolation–ensemble Kalman filter for the Canadian Precipitation Analysis, Nonlinear Processes in Geophysics, 29, 329–344, https://doi.org/10.5194/npg-29-329-2022, 2022.

Khedhaouiria, D., Gasset, N., Fortin, V., Lauer, A., Dimitrijevic, M., Bulat, M., and Roy, G.: CaSR V3: ECCC hourly 0.1 degrees surface reanalysis across North America, Earth System Science Data Discussion, 2025.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., and Kirschbaum, D. B.: So, How Much of the Earth's Surface Is Covered by Rain Gauges?, Bulletin of the American Meteorological Society, 98, 69–78, https://doi.org/10.1175/BAMS-D-14-00283.1, 2017.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, Journal of Hydrology, 424-425, 264–277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrology and Earth System Sciences, 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: General Specifications and Basic Characteristics, Journal of the Meteorological Society of Japan. Ser. II, 93, 5–48, https://doi.org/10.2151/jmsj.2015-001, 2015.

L'Ecuyer, T. S. and Stephens, G. L.: An Estimation-Based Precipitation Retrieval Algorithm for Attenuating Radars, Journal of Applied Meteorology, 41, 272–285, https://doi.org/10.1175/1520-0450(2002)041<0272:AEBPRA>2.0.CO;2, 2002.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

Lespinas, F., Fortin, V., Roy, G., Rasmussen, P., and Stadnyk, T.: Performance Evaluation of the Canadian Precipitation Analysis (CaPA), Journal of Hydrometeorology, 16, 2045–2064, https://doi.org/10.1175/JHM-D-14-0191.1, 2015.

Lespinas, F., Roy, G., and Mahidjiba, A.: Implementation of the Regional Deterministic Precipitation Analysis System (CaPA-RDPA) version 6.0.0, Technical Note, Meteorological Service of Canada and Meteorological Research Division, Environment and Climate Change Canada, https://collaboration.cmc.ec.gc.ca/cmc/cmoi/product_guide/docs/lib/technote_capa_rdpa-600_e.pdf, 2024.

Lopez, P.: Direct 4D-Var Assimilation of NCEP Stage IV Radar and Gauge Precipitation Data at ECMWF, Monthly Weather Review, 139, 2098–2116, https://doi.org/10.1175/2010MWR3565.1, 2011.

Mahfouf, J. F., Brasnett, B., and Gagnon, S.: A Canadian precipitation analysis (CaPA) project: Description and preliminary results, Atmosphere - Ocean, 45, 1–17, https://doi.org/10.3137/ao.v450101, 2007.

McTaggart-Cowan, R., Vaillancourt, P. A., Zadra, A., Chamberland, S., Charron, M., Corvec, S., Milbrandt, J. A., Paquin-Ricard, D., Patoine, A., Roch, M., Separovic, L., and Yang, J.: Modernization of Atmospheric Physics Parameterization in Canadian NWP, Journal of Advances in Modeling Earth Systems, 11, 3593–3635, https://doi.org/10.1029/2019MS001781, 2019.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J. N.: ERA5-Land: A state-of-the-art global reanalysis dataset for land applications, Earth System Science Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, publisher: Copernicus GmbH, 2021.

Nykanen, D. K., Foufoula-Georgiou, E., and Lapenta, W. M.: Impact of Small-Scale Rainfall Variability on Larger-Scale Spatial Organization of Land–Atmosphere Fluxes, Journal of Hydrometeorology, 2, 105–121, https://doi.org/10.1175/1525-7541(2001)002<0105:IOSSRV>2.0.CO;2, 2001.

Panthou, G., Vischel, T., Lebel, T., Blanchet, J., Quantin, G., and Ali, A.: Extreme rainfall in West Africa: A regional modeling, Water Resources Research, 48, 1–19, https://doi.org/10.1029/2012WR012052, 2012.

Parker, W. S.: Reanalyses and observations: What's the Difference?, Bulletin of the American Meteorological Society, 97, 1565–1572, https://doi.org/10.1175/BAMS-D-14-00226.1, publisher: American Meteorological Society, 2016.

Pradhan, R. K., Markonis, Y., Marra, F., Nikolopoulos, E. I., Papalexiou, S. M., and Levizzani, V.: Diurnal variability of global precipitation: insights from hourly satellite and reanalysis datasets, Hydrology and Earth System Sciences, 29, 4929–4949, https://doi.org/10.5194/hess-29-4929-2025, 2025.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Van Den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP Climate Forecast System Reanalysis, Bulletin of the American Meteorological Society, 91, 1015–1058, https://doi.org/10.1175/2010BAMS3001.1, 2010.

Smith, A., Lott, N., and Vose, R.: The Integrated Surface Database: Recent Developments and Partnerships, Bulletin of the American Meteorological Society, 92, 704–708, https://doi.org/10.1175/2011BAMS3015.1, 2011.

Smith, C. D., Mekis, E., Hartwell, M., and Ross, A.: The hourly wind-bias-adjusted precipitation data set from the Environment and Climate Change Canada automated surface observation network (2001–2019), Earth System Science Data, 14, 5253–5265, https://doi.org/10.5194/essd-14-5253-2022, 2022.

https://doi.org/10.5194/egusphere-2026-620
Preprint. Discussion started: 17 April 2026

Soci, C., Bazile, E., Besson, F. O., and Landelius, T.: High-resolution precipitation re-analysis system for climatological purposes, Tellus, Series A: Dynamic Meteorology and Oceanography, 68, https://doi.org/10.3402/tellusa.v68.29879, publisher: Co-Action Publishing, 2016.

Verrelle, A., Glinton, M., Bazile, E., Le Moigne, P., Randriamampianina, R., Ridal, M., Berggren, L., Undén, P., Schimanke, S., Mladek, R., and Soci, C.: CERRA-Land sub-daily regional reanalysis data for Europe from 1984 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.a7f3cd0b, 2022.

Wang, X. L., Xu, H., Qian, B., Feng, Y., and Mekis, E.: Adjusted Daily Rainfall and Snowfall Data for Canada, Atmosphere-Ocean, 55, 155–168, https://doi.org/10.1080/07055900.2017.1342163, 2017.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M.: The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment, Ocean Science, 15, 779–808, https://doi.org/10.5194/os-15-779-2019, 2019.